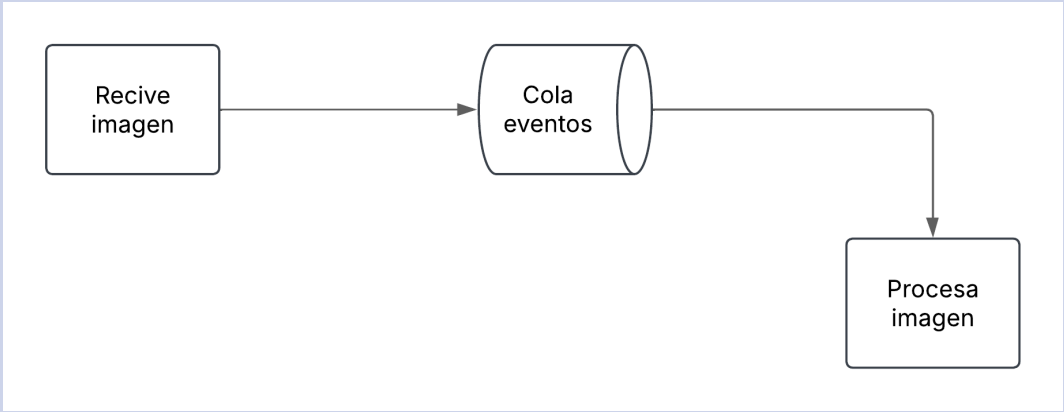


# MISO

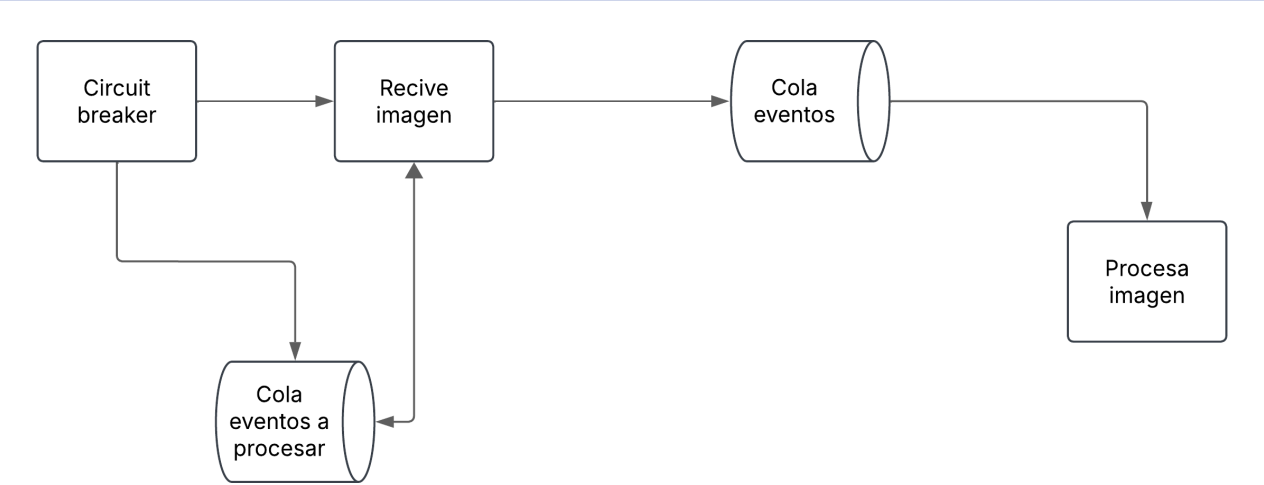
Maestría en Ingeniería de Software

## Entrega 3: Diseño de la experimentación y POC de servicio

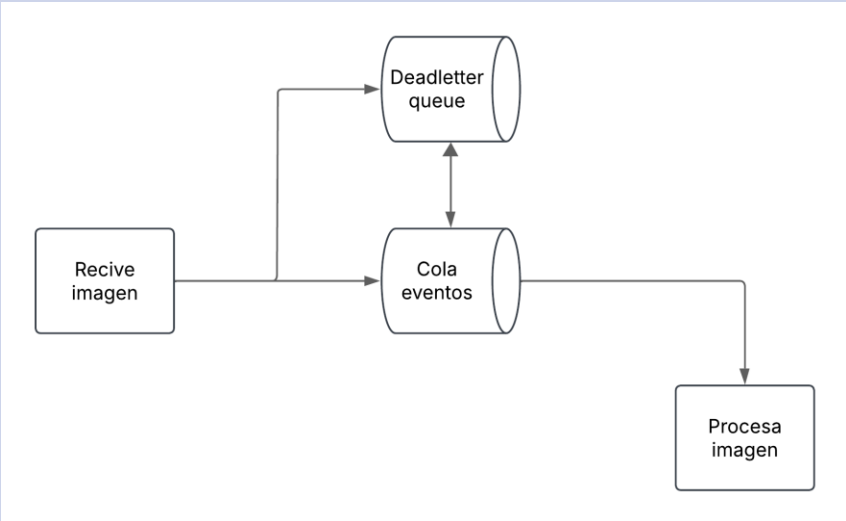
# Atributo de calidad 1: Escalabilidad

Escenario de calidad: Escalabilidad de lectura y escritura			
Escenario #: 1	Durante un pico de carga de datos al sistema, este debe ser capaz de soportarlo sin afectar la integridad del sistema como un todo		
Fuente	Usuario		
Estímulo	Carga de datos		
Ambiente	Producción con alta demanda		
Artefacto	Gestor de archivos		
Respuesta	El sistema es capaz de aceptar las nuevas imágenes		
Medida de la respuesta	El sistema responde en menos de 5 segundos		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Arquitectura basada en eventos	Publica en tópico de comandos y recibe respuesta en el tópico de eventos	Incrementa la complejidad del sistema para ser mantenido a futuro	Genera consultas bloqueantes
CQRS	Aislamiento de lectura y escritura	Mayor complejidad de implementación y entendimiento en el código	Bloquear la base de datos por locks
Justificación	Por medio de la comunicación basada en eventos podemos asegurar que ninguna micro servicio se comunica de punto a punto, asegurando así que no se bloquean los servicios. Adicionalmente con CQRS aseguramos la separación de escritura y lectura, para el caso en que alguien consulte las imágenes al tiempo que se escribe no genere mucha sobre carga		
Diagrama de arquitectura			

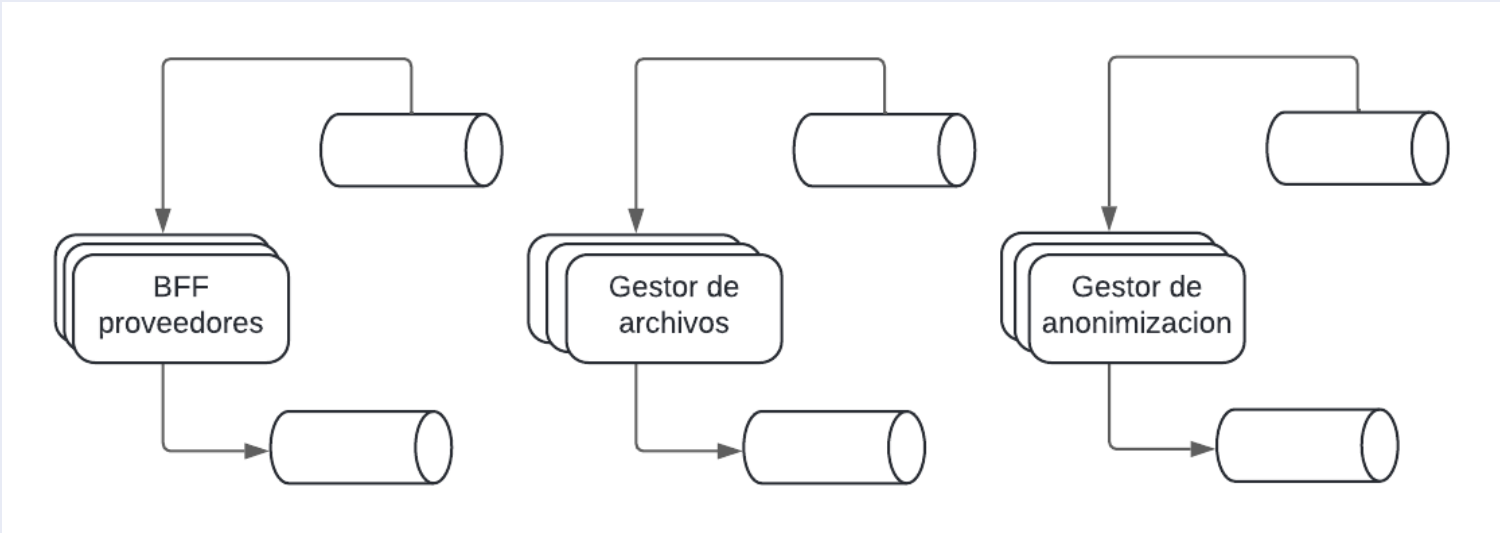
# Atributo de calidad 1: Escalabilidad

Escenario de calidad: Escalabilidad ante alto consumo de CPU			
Escenario #: 2	Durante un sobrecargo de CPU debido al procesamiento de anonimización de imágenes		
Fuente	El microservicio de anonimización y el usuario que manda las imágenes		
Estímulo	Hay bastantes imágenes siendo procesadas por una instancia		
Ambiente	Producción, con alta demanda de imágenes por procesar		
Artefacto	Anonimizador		
Respuesta	El sistema logra escalar para atender la demanda de las imágenes		
Medida de la respuesta	Escala en menos de 30 segundos		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Procesamiento de eventos	Comunicación asincronica con el topico que recibe el procesameitno de una imagen	Agrega una mayor complejidad al momento de implementar la solución en código	Puede generar transacciones bloqueantes
Circuit Breaker	Maneja la posibilidad seguir procesando peticiones o detener el procesamiento	Agrega nivel de complejidad mayor en el código y en el entendimeinto del sistema como un todo	Puede dar de baja un sistema si no se implementa de forma correcta
Justificación	Con el circuit breaker podemos manejar el estado del microservicio, es decir, si se encuentra sobre cargado y empieza a responder con error, el microservicio se puede cerrar hasta que se estabilice y todos los eventos son encolados hasta que se recupere el microservicio, es decir, hasta que haya escalado para poder asumir la carga que está recibiendo		
Diagrama de arquitectura	 <pre>graph LR; CB[Circuit breaker] --&gt; RI[Recive imagen]; CB --&gt; CEP[Cola eventos a procesar]; RI --&gt; CE[Cola eventos]; CE --&gt; PI[Procesa imagen]; CEP --&gt; RI;</pre>		

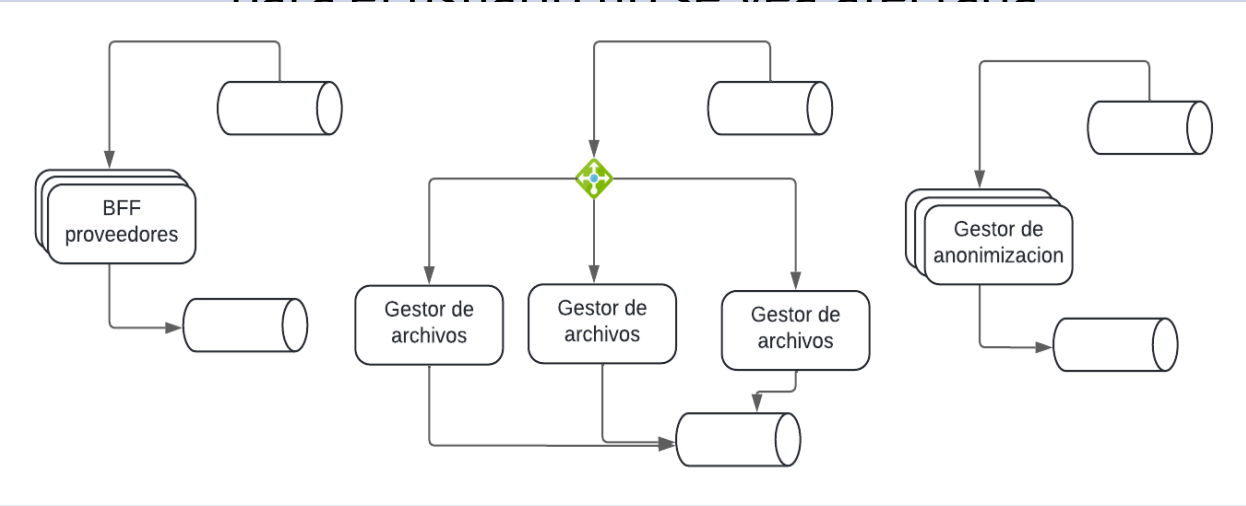
# Atributo de calidad 1: Escalabilidad

Escenario de calidad: Escalabilidad ante una recuperación de servicio con alta cantidad de eventos sin procesar			
Escenario #: 3	Después de una caída del microservicio de procesamiento de imágenes los eventos encolados son procesados pero generan un alto consumo de memoria y cpu		
Fuente	Cola de eventos de procesamiento de imágenes		
Estímulo	Hay bastantes eventos encolados		
Ambiente	Producción normal, después de recuperarse de una caída y con volumen alto de eventos encolados		
Artefacto	Gestor de archivos		
Respuesta	El microservicio es capaz de escalar para procesar los eventos encolados y después escala a su cantidad normal de operación		
Medida de la respuesta	No debe tomar más de 30 segundos		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Deadletter queues	Permite centralizar en un solo lugar los eventos fallidos	Agrega una mayor complejidad al sistema y el entendimiento del sistema como un todo	
Políticas de autoescalamiento	Permite empezar a escalar después de cierto uso de memoria y cpu	Mayor complejidad en la configuración del microservicio	
Justificación	La deadletter queues nos permitirá almacenar todos los eventos fallidos y recuperarlos de esa cola sin afectar las demás colas. Adicionalmente, las políticas de auto escalamiento nos permitirán escalar el sistema después de una falla si hay bastante número de eventos encolados y la CPU y memoria empiezan a consumirse		
Diagrama de arquitectura			

# Atributo de calidad 2: Disponibilidad

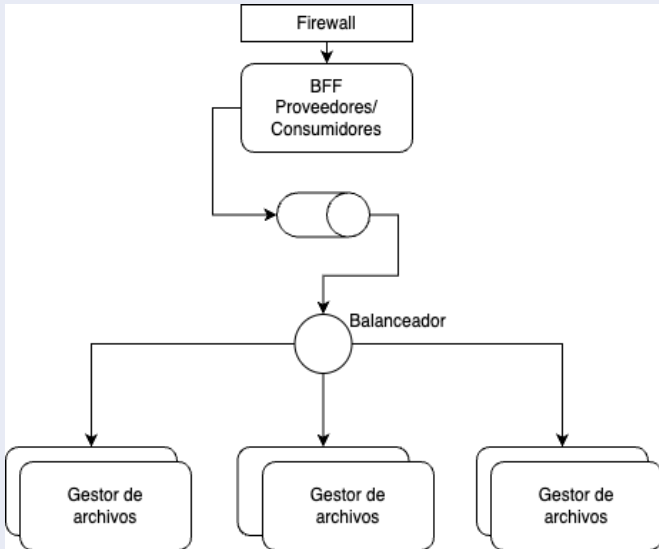
Escenario de calidad: Disponibilidad del sistema de procesamiento de imágenes médicas durante picos de carga			
Escenario #: 4	El sistema debe garantizar una alta disponibilidad incluso durante picos de carga, como cuando múltiples centros de salud envían grandes volúmenes de imágenes médicas simultáneamente.		
Fuente	Centros de salud y laboratorios		
Estímulo	Varios centros de salud envían grandes volúmenes de imágenes médicas al mismo tiempo.		
Ambiente	Operación normal, con un alto volumen de datos siendo procesados simultáneamente por múltiples centros de salud en diferentes zonas horarias.		
Artefacto	Sistema de procesamiento de imágenes médicas de STA, incluyendo los componentes de anonimización.		
Respuesta	El sistema debe procesar y almacenar las imágenes médicas de manera eficiente, garantizando que no haya interrupciones en el servicio y que los datos estén disponibles.		
Medida de la respuesta	El sistema debe mantener una disponibilidad del 99.9% por año.		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Implementar un sistema de comunicación basada en eventos	El sistema debe mantener una disponibilidad del 99.9% por año	Implementar un sistema basado en eventos y microservicios mejora la disponibilidad, pero incrementa la complejidad en la gestión, monitoreo y mantenimiento de la arquitectura.	Si el sistema de eventos no está correctamente configurado o monitoreado, podría haber interrupciones en el servicio o pérdida de datos.
Microservicios			
Justificación	El enfoque basado en eventos, como arquitectura de software, optimiza la disponibilidad del sistema al distribuir dinámicamente las tareas de procesamiento entre múltiples servidores. Esto permite que el servicio permanezca operativo incluso cuando varios centros de salud envíen grandes volúmenes de datos simultáneamente.		
Diagrama de arquitectura			

# Atributo de calidad 2: Disponibilidad

Escenario de calidad: Disponibilidad del sistema de almacenamiento de datos médicos durante fallos de hardware			
Escenario #: 5	El sistema debe garantizar que los datos médicos estén disponibles con una interrupción máxima de 30 segundos en caso de fallos de hardware en los servidores de almacenamiento.		
Fuente	Desarrolladores de IA y centros de salud		
Estímulo	Un fallo de hardware en uno de los servidores de almacenamiento que contiene datos médicos.		
Ambiente	Operación normal, con la posibilidad de fallos de hardware en los servidores de almacenamiento.		
Artefacto	Sistema de almacenamiento de archivos.		
Respuesta	El sistema debe garantizar que los datos médicos sigan estando disponibles con una interrupción máxima de 30 segundos en caso de fallos de hardware, mediante la redundancia de datos.		
Medida de la respuesta	El sistema debe garantizar que la interrupción del servicio no supere los 30 segundos durante un fallo de hardware.		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Implementar un sistema de almacenamiento con replicación de datos en múltiples servidores.	El sistema debe garantizar que la interrupción del servicio no supere los 30 segundos durante un fallo de hardware.	Mayor complejidad en la gestión de la infraestructura	Si el sistema de replicación y conmutación por error no funciona correctamente, lo que llevaría a interrupciones o caídas de servicios completos
Implementar un sistema que detecte y redirija las solicitudes a servidores redundantes		Costos más altos debido a la necesidad de redundancia	
Justificación	Al contar con la información replicada, en caso de que uno de los servidores experimente un fallo, el balanceador de carga podrá redirigir las solicitudes a un servidor en funcionamiento, garantizando que la disponibilidad de la información para el usuario no se vea afectada.		
Diagrama de arquitectura			



# Atributo de calidad 2: Disponibilidad

Escenario de calidad: Disponibilidad del sistema al consumo de imágenes médicas anonimizadas frente a un ataque de denegación de servicio (DDoS)			
Escenario #: 6	El sistema debe garantizar la disponibilidad del servicio de acceso a imágenes médicas anonimizadas en caso de un ataque DDoS mitigando el impacto sin afectar a los usuarios legítimos.		
Fuente	Atacante		
Estímulo	Sobrecargar en la infraestructura de STA debido a una cantidad masiva de solicitudes simultáneas a los servicios definidos.		
Ambiente	Operación normal, con el sistema desplegado en múltiples zonas de disponibilidad.		
Artefacto	Sistema de almacenamiento de archivos.		
Respuesta	El sistema debe detectar el ataque DDoS, mitigar el tráfico malicioso y garantizar el acceso continuo a los servicios de acceso de imágenes médicas anonimizadas a los usuarios legítimos sin degradar el rendimiento.		
Medida de la respuesta	El sistema debe mantener la tasa de disponibilidad mantenida en 99.99% durante el ataque DDoS.		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Implementación de CDN y balanceadores de carga para distribuir el tráfico.	La capacidad de detección rápida y bloqueo de tráfico malicioso sin afectar a usuarios legítimos.	Implementación de protección avanzada contra DDoS puede aumentar la latencia mínima.	Si no se mitiga correctamente, el sistema puede quedar indisponible por un tiempo prolongado afectando a los usuarios.
Autoescalado de servidores para manejar cargas variadas.			La degradación en la experiencia del usuario en los servicios de acceso a imágenes médicas anonimizadas.
Justificación	Al incluir un firewall y un balanceador de carga entre el BFF Proveedores/Consumidores y Gestor de archivos permite que frente a un ataque de DDoS el sistema de almacenamiento de archivos esté disponible en todo momento a los usuarios.		
Diagrama de arquitectura			

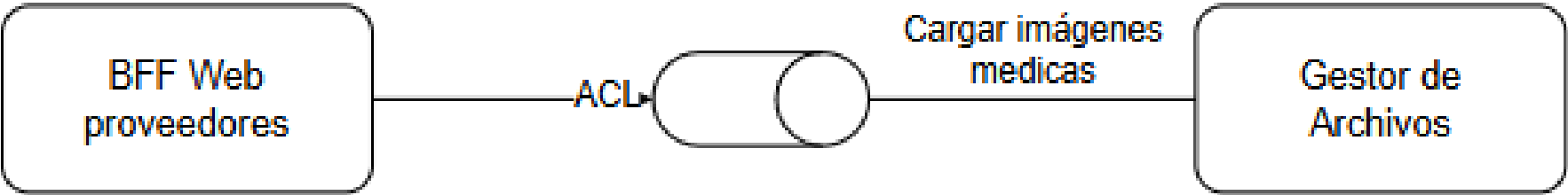
# Atributo de calidad 3: Configurabilidad/Extensibilidad

Escenario de calidad: Configurabilidad para las diferentes regulaciones y mercados en términos de procesamiento de imágenes médicas y diagnósticos			
Escenario #: 7	El sistema debe ser altamente configurable para adaptarse a las diferentes regulaciones y requisitos en términos de privacidad de datos y procesamiento de anonimizado.		
Fuente	Reguladores de cada país (HIPAA en USA, GDPR en Europa y otros)		
Estímulo	La introducción de nuevas leyes o las actualizaciones en las normativas de privacidad de datos.		
Ambiente	Operación normal, el sistema operando en diferentes regiones con regulaciones específicas.		
Artefacto	Sistema de configuración y procesamiento de imágenes médicas y diagnósticos		
Respuesta	El sistema debe poder adaptarse a las nuevas regulaciones en el menor tiempo posible cumpliendo con las regulaciones y normativas proporcionadas.		
Medida de la respuesta	En menos de 24 horas el sistema de poder adaptarse a los cambios requeridos por los reguladores.		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Diseñar el sistema con módulos independientes que puedan ser configurados según las regulaciones.	Un sistema altamente configurable es más difícil de mantener y la complejidad aumenta.	Aumentar la configurabilidad puede incrementar la complejidad del sistema.	Una configuración incorrecta podría resultar en incumplimiento de las regulaciones y normativas.
Implementar un sistema de gestión de cambios que permita actualizar las configuraciones rápidamente.		Un sistema altamente configurable puede tener un impacto en el rendimiento.	
Justificación	Al tener un módulo transversal encargado de la configurabilidad de las regulaciones y mercados se logra que frente a cualquier cambio requerido por los reguladores los módulos del sistema puedan adaptarse de forma ágil y concreta.		
Diagrama de arquitectura	<div>Configurabilidad</div> <pre>graph LR; ADP1[ADP] --&gt; GA[Gestor de archivos]; GA --&gt; S1(( )); S1 --&gt; AN[Anonizador]; AN -- ADP --&gt; S2(( )); S2 --&gt; VA[Validador de anonimización];</pre> <p>El diagrama de flujo, titulado 'Configurabilidad', muestra un proceso de flujo de datos. Comienza con un nodo de entrada etiquetado como 'ADP' que se conecta a un rectángulo 'Gestor de archivos'. Desde el 'Gestor de archivos', la línea de flujo continúa a un primer símbolo de almacenamiento (un rectángulo con dos semicírculos opuestos). Desde este primer almacenamiento, la línea de flujo se dirige a un rectángulo 'Anonizador'. Desde el 'Anonizador', la línea de flujo pasa por un segundo símbolo de almacenamiento (otro rectángulo con dos semicírculos opuestos) y finalmente se conecta a un rectángulo 'Validador de anonimización'. Hay una etiqueta 'ADP' entre el 'Anonizador' y el segundo símbolo de almacenamiento.</p>		

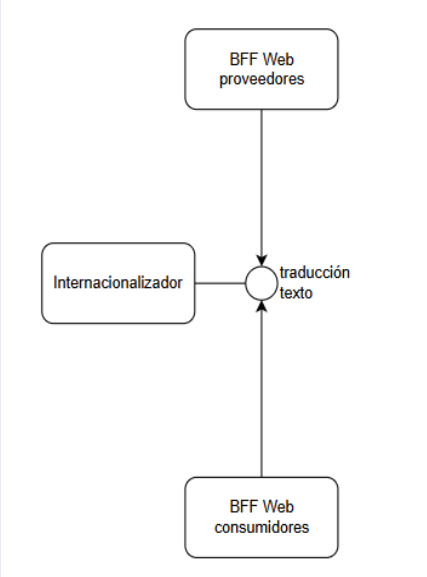


# Atributo de calidad 3: Configurabilidad/Extensibilidad

Escenario de calidad: Almacenamiento dinámico

Escenario #: 8	El sistema en un ambiente normal debe persistir los archivos cargados cumpliendo con la normativa de cada pais		
Fuente	Usuario		
Estímulo	Proveedor carga archivo		
Ambiente	Operación normal		
Artefacto	Gestor de archivos		
Respuesta	El gestor debe persistir el archivo en una base de datos del país desde el que se hace el consumo		
Medida de la respuesta	El archivo cargado debe ser almacenado en una base de datos del país de consumo en máximo 20 segundos		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Cargado de archivo mediante colas de mensajería	Los archivos deben ser persistidos en un tiempo máximo de 20 segundos	El tener múltiples bases de datos añade complejidad en el mantenimiento y soporte del aplicativo	Pico de mensajes en la cola que provoque demora en el proceso de carga de imágenes medicas
Variables de entorno para definir la persistencia para cada país		Hay países que requieren mayor seguridad en el proceso de almacenado de información medica	Incumplimiento de normativas de algún país en la seguridad de las bases de datos
Justificación	El uso de colas evita el acoplamiento con las bases de datos o los proveedores del servicio, haciendo que sea más fácil hacer cambios si se es necesario. Las variables de entorno facilitan el cambio en la configuración del servicio en caso de necesitarse un cambio.		
Diagrama de arquitectura			

# Atributo de calidad 3: Configurabilidad/Extensibilidad

Escenario de calidad: Internacionalizador			
Escenario #: 9	El sistema en un ambiente normal debe adaptarse al idioma del país de consumo		
Fuente	Usuario		
Estímulo	Acceso al Frontend desde Estados unidos o Latinoamérica		
Ambiente	Operación normal		
Artefacto	Frontend		
Respuesta	El Frontend debe adaptarse al idioma del país de consumo		
Medida de la respuesta	La pantalla debe cargar en máximo 2 segundos en cargar en el idioma correspondiente		
Decisiones Arquitecturales	Punto de sensibilidad	Tradeoff	Riesgo
Modulo internacionalizador de texto para los Frontends	El proceso de internacionalización debe tener baja latencia para cumplir con el tiempo máximo de 2 segundos de carga de pantalla	Tener un frontend flexible que se ajusta a cada región puede aumentar los tiempos de carga de las pantallas	Errores de traducción
Detección automática de región por parte de los Frontends			Alta latencia con el módulo internacionalizador
Justificación	El tener un módulo encargado de internacionalizar favorece la armonía entre los Frontends. La detección automática de región favorece a la solución para dar el servicio esperado dependiendo del país.		
Diagrama de arquitectura	 <pre>graph TD; A[BFF Web proveedores] --&gt; C((traducción texto)); B[Internacionalizador] --&gt; C; D[BFF Web consumidores] --&gt; C;</pre>		