Universidad de La Habana
Facultad de Matemática y Computación

Sistema de Autenticación Central

Autor:
Nadia González Fernández

Tutores:
Lic. Roberto Martí Cedeño

Trabajo de Diploma
presentado en opción al título de
Licenciado en Ciencia de la Computación

Noviembre de 2022

github.com/nala7/central_authentication_system

Dedicación

Agradecimientos

Agradecimientos

Opinión del tutor

Opiniones de los tutores

Resumen

Resumen en español

Abstract

Resumen en inglés

Índice general

Índice de figuras

Ejemplos de código

Introducción

Hoy en día las personas usan muchos sistemas de software independientes con distintas identidades. La identidad del usuario es considerado un conjunto permanente o de larga vida de atributos asociados a la identidad de un usuario. Para obtener acceso a servicios muchos requieren que el usuario esté registrado y haya iniciado sesión. Muchas de las veces los usuarios están registrados en los distintos sitios con el mismo nombre de usuario y con la misma o similar contraseña, lo cual no es la mejor práctica para preservar la seguridad de las cuentas. Por otra parte, es fácil de olvidar las credenciales, lo cual obliga a los gestores del sistema a enviar un correo no encriptado con información confidencial.

Por ello el manejo de múltiples nombres de usuarios y contraseña es una tarea molesta del Internet actual.

Teniendo en cuenta todos esos problemas de seguridad, cada vez más administradores de sitios web deciden delegar esos servicios de administración y autenticación a terceros, entidades externas especializadas en ese tipo de actividad. Pueden alojar, almacenar, administrar y proteger los datos de los usuarios de un proveedor de servicios en particular. Ofrecen una interfaz de programación de aplicaciones (conocida también por la sigla API, en inglés, *Application Programming Interface*), que pueden proporcionar acceso a datos de usuario para aplicaciones externas. Estas soluciones permiten compartir los datos con más de un sistema web, permitiendo así el Inicio de Sesión Único (SSO) para aplicaciones de terceros.[1]

El Nodo Central de la Universidad de La Habana tiene entre sus responsabilidades dar las credenciales digitales a todos los usuarios de la Universidad. Los trabajadores y estudiantes de la institución registran sus datos personales en las bases de datos de recursos humanos y secretaria docente respectivamente. Esa información es utilizada más adelante para que los usuarios puedan autenticarse en los distintos servicios de la institución, respaldado por sus sistemas de origen.

En este trabajo se presenta una solución a los problemas de autenticación que se ajusta a las necesidades y posibilidades de la Universidad de La Habana.

Motivación

Actualmente vivimos en un mundo donde la tecnología tiene un papel protagónico. La Universidad de La Habana en los últimos años ha estado inmersa en el proceso de transformación digital que lleva a cabo nuestro país, como continuidad de la estrategia de informatización de la sociedad cubana, que pretende integrar las tecnologías digitales a todos los ámbitos de la sociedad.

En este sentido, nuestra sede universitaria ha avanzado en la digitalización de la información y los procesos y trabaja para facilitar el acceso a la red de todos sus estudiantes y trabajadores. A la par de estos avances, debe crecer también la seguridad de los sistemas para evitar brechas de información y asegurar la privacidad de los usuarios.

Garantizar un acceso seguro y sencillo a los recursos de la red es esencial para alcanzar este objetivo. Por consiguiente, se deben establecer mecanismos que cuenten con un alto nivel de seguridad y permitan identificar quién realmente está autorizado para acceder a los recursos del sistema.

La creación de una plataforma central de autenticación resolvería muchos de estos problemas y mejoraría la experiencia de los usuarios. Estos solo tendrían una cuenta para acceder a todos los servicios que ofrece la Universidad, por lo que no tendrían que lidiar con formularios de inicio de sesión cada vez que vayan a acceder a un servicio.

Antecedentes

Actualmente todos los usuarios de la Universidad de La Habana se almacenan en dos Protocolos Ligeros de Acceso a Directorios (en inglés: Lightweight Directory Access Protocol, también conocido por sus siglas como LDAP). En los dos directorios se guardan las cuentas de correo y los datos de sus usuarios. En uno ellos se registra la información de los estudiantes y en el otro la de los trabajadores. A partir de estos, todos los sitios web y aplicaciones de la universidad, individualmente, verifican la pertenencia del usuario a la institución.

Cada aplicación y servicio tiene un servicio de autenticación individual que dependen de los dos mencionados directorios que contienen a los usuarios.

Problemática

La Universidad brinda servicios como Wi-Fi, correo, proxy y EVEA (Entorno Virtual de Enseñanza Aprendizaje), imprescindibles para el funcionamiento de la institución. También brinda servicios más especializados como los sistemas de contabilidad, recursos humanos, inventarios, el registro de notas, entre otros. Todos estos

servicios utilizan la autenticación de forma individual, lo cual significa que un usuario tiene diversas cuentas a pesar de pertenecer todas a la misma institución.

El sistema implementado actualmente es poco eficiente y requiere de intervención humana constante para corregir y/o restablecer el apropiado funcionamiento de los servicios de autenticación. Esto genera tiempos elevados de respuesta y dificulta el mencionado proceso de transformación digital que está siendo llevado a cabo por nuestra Universidad.

Objetivo

Con el propósito de presentar una propuesta para solucionar la problemática expuesta anteriormente, se plantean los siguientes objetivos:

Objetivo General

- Diseñar e implementar un sistema de autenticación centralizada para todos los usuarios de La Universidad de La Habana.

Objetivos Específicos

- Generar los servicios de autenticación con compatibilidad con todas las tecnologías existentes y previstas en la institución.
- Implementar la gestión de control de acceso a todos los servicios ofrecidos por el Nodo Central de forma extensible a futuros servicios y fuentes de datos.

Estructura de la Tesis

Capítulo 1

Estado del Arte

En este capítulo se brindan las definiciones de herramientas utilizadas para la autenticacióna. También se realiza un estudio sobre el estado del arte de las mismas. Además, se brindan razones para incluir su utilización como parte de la solución propuesta.

1.1. Inicio de Sesión Único

El Inicio de Sesión Único (en inglés *Single Sign-On* o también conocido por sus siglas SSO) ha sido ampliamente adoptado para la autenticación en línea debido a su utilidad y seguridad que ofrece. Este es un método de autenticación que permite a los usuarios iniciar sesión con un único conjunto de credenciales en varios sistemas de software independientes. El inicio de sesión único facilita que un usuario no tenga que iniciar sesión en cada aplicación que use. Con este servicio los usuarios pueden acceder a todas las aplicaciones necesarias sin tener que autenticarse con otras credenciales.[2]

En el mundo digital actual, los usuarios acceden a múltiples sistemas para llevar a cabo sus quehaceres. A medida que aumenta la cantidad de sistemas, también aumenta la cantidad de credenciales de cada usuario y, por lo tanto, también incrementa la posibilidad de perderlas u olvidarlas. El inicio de sesión único se puede utilizar para resolver muchos problemas relacionados con múltiples credenciales para diferentes aplicaciones. El acceso de inicio de sesión único al centro de autenticación principal permite a los usuarios obtener acceso a todos los demás recursos disponibles. SSO ayuda a mejorar la productividad del usuario y del desarrollador al evitar que el usuario recuerde varias contraseñas y también reduce la cantidad de tiempo que el usuario dedica a escribir varias contraseñas para iniciar sesión. SSO también simplifica la administración mediante la gestión de credenciales únicas en lugar de múltiples credenciales. Facilita la gestión de los derechos de un usuario que llega, cambia de función dentro o sale de la empresa, para integrar rápidamente aplicaciones adicionales,

delegar derechos de acceso durante las vacaciones sin aumentar la carga de trabajo de la mesa de ayuda. [3]

Grandes empresas como Google, Facebook y Microsoft utilizan estos servicios. En particular Google permite a sus usuarios iniciar sesión una única vez para tener acceso a todos los servicios de la empresa que se encuentran en la nube. Cuando se configura SSO, los usuarios pueden iniciar sesión en terceros proveedores de identidad y luego acceder a las aplicaciones de Google directamente sin un segundo inicio de sesión [4]. Por ejemplo, si se accede a un servicio de Google como Gmail, se autentica automáticamente en YouTube, AdSense, Google Analytics, y otras applicaciones de Google. Del mismo modo, si cierra la sesión de su Gmail u otras aplicaciones de Google, se cerrará automáticamente la sesión de todas las aplicaciones; esto se conoce como Cierre de Sesión Único (en inglés: *Single Logout*) [5]

[3]

1.2. OpenID Connect

OpenID Connect (también OIDC)es una capa de identidad simple implementada a partir del protocolo *OAuth 2.0*. Permite a los clientes verificar la identidad del usuario final en función de la autenticación realizada por un servidor de autorización, así como obtener información básica del perfil del Usuario final de manera interoperable y similar al protocolo REST. [6].

OpenID Connect es uno de los protocolos de tipo *Single Sign-On* más utilizados para delegar la autenticación. También tiene un formato simple, por lo que ha ganado popularidad y es soportado por grandes empresas como Google, IBM, Microsoft, Amazon y PayPal [7]. La nueva versión es compatible con *API* y puede ser usado por aplicaciones nativas y móviles. También define mecanismos opcionales más robustos para firmas y cifrados. [6]

Este protocolo está basado en *OAuth 2.0* por lo que tiene todas las ventajas de este protocolo. Sin embargo, lo extiende ya que ofrece facilidades para obtener más información de la identidad de los usuarios eficientemente. Permite un flujo de información adicional que genera un *id-token* que contiene datos del usuario. De esta forma las aplicaciones no solo tienen acceso a los permisos de los usuarios, sino también obtienen información sobre la identidad de los usuarios. [6][1]

1.3. Otros protocolos de autenticación

1.3.1. SAML

Lenguaje de Marcado para Confirmaciones de Seguridad, conocido como SAML (en inglés: *The Security Assertion Markup Language*) es un marco de trabajo que permite expresar asertos[1] acerca de la identidad, los atributos y las autorizaciones de un sujeto con el objetivo de facilitar las relaciones entre distintas empresas, así como las relaciones de estas con sus usuarios. Este marco de trabajo permite a las compañías crear identidades federadas, lo cual les facilita las tareas de gestión de perfiles, autenticación y autorización de usuarios. El caso típico de uso es el de *Single-Sign-On* (SSO), que permite a los usuarios acceder a diversos sitios en la federación con una única autenticación. [8]

SAML y OpenID Connect son protocolos de identificación, diseñados para autenticar a los usuarios. También proporcionan datos de identidad para el control de acceso y como método de comunicación para la identidad de un usuario.

SAML durante muchos años ha proporcionado un medio seguro de intercambio de datos de identidad, por lo que muchas organizaciones confían en él. También es muy rico en funciones y cubre una amplia gama de requisitos de identidad.

OIDC, al ser más nuevo y en se encuentra desarrollo, todavía está rezagado con respecto a SAML en términos de características. Sin embargo, para muchas aplicaciones donde solo se necesita un requisito simple para los datos de identidad básicos, particularmente en el espacio del consumidor, OIDC es muy útil, ya que es mucho más fácil de usar que SAML y no requiere el manejo pesado de XML que requiere SAML.

Actualmente, SAML se usa prinipalmente para la identificacion de ciudadanos del gobierno y la autenticación empresarial. Sin embargo, esto está comenzando a cambiar, con sistemas más modernos que utilizan OIDC en lugar de SAML. Esto se debe a que OIDC permite un procesamiento de datos más liviano que SAML, utilizando tokens JSON (token de ID) en lugar de XML. OIDC es ideal para usar con aplicaciones móviles y aplicaciones web de una sola página, donde el uso de SAML sería complicado. [9]

OIDC se adapta a las condiciones del Nodo Central de la Universidad de La Habana ya que se cuentan con pocos recursos, por lo que es necesario un procesamiento ligero de información.

[1] Los asertos definen las afirmaciones de seguridad de una entidad dentro de un sistema. Estas afirmaciones pueden ser de tres tipos: asertos de autenticación, de atributos y de decisiones de autorización.

1.4. LDAP

El Protocolo Ligero de Acceso a Directorios (en inglés: *Lightweight Directory Access Protocol*, también conocido por sus siglas de LDAP) es un conjunto de protocolos de licencia abierta que son utilizados para acceder a la información que está almacenada de forma centralizada en una red. Este protocolo se utiliza a nivel de aplicación para acceder a los servicios de directorio remoto. [10]

LDAP está basado en estándares implementados sobre TCP/IP. Permite a los clientes interactuar directamente con los servidores de los directorios: almacenar y consultar información, buscar datos filtrados, autenticar usuarios, entre otros.

Este protocolo es utilizado actualmente por muchos sistemas que apuestan por el software libre al utilizar distribuciones de Linux para ejercer las funciones propias de un directorio activo en el que se gestionarán las credenciales y permisos de los usuarios y estaciones de trabajo en redes LAN corporativas en conexiones cliente/servidor mientras que open id connect y el flujo de emisión de tokens esta pensado para redes abiertas (internet).

Un directorio remoto es un conjunto de objetos que están organizados de forma jerárquica, tales como: nombre, claves, direcciones, etc. Estos objetos estarán disponibles para una serie de clientes conectados mediante una red, normalmente interna o LAN, y proporcionarán las identidades y permisos para esos usuarios que la utilicen.

Sería bueno que a cada protocolo le incluyas una url con el rfc que lo define (cualquiera de ellos, el primero o el último) (puede haber mas de un rfc por protocolo)
Ahora si esta en tu objetivo de tesis entender un protocolo, su link lo pones en la bibliografía consultada, por el contrario si es adicional, ponlo como un pie de página.

LDAP está basado en el protocolo X.500 para compartir directorios, y contiene esta información de forma jerarquizada y mediante categorías para proporcionarnos una estructura intuitiva desde el punto de vista de la gestión por parte de los administradores.

Estos directorios se utilizan generalmente para contener información virtual de usuarios, para que otros usuarios accedan y dispongan de información acerca de los contactos que están aquí almacenados. Además es capaz de comunicarse de forma remota con otros directorios LDAP situados en servidores que pueden estar en el otro lado del mundo para acceder a la información disponible. De esta forma se crea una base de datos de información descentralizada y completamente accesible.

El sistema de autenticación vigente en el Nodo Central verifica sus usuarios con dos sistemas implementados con LDAP. Este protocolo se adapta a las necesidades y condiciones actuales de la Universidad ya que es *Open Source* (OSS o código abierto).

1.5. Active Directory

Directorio Activo (en inglés: Active Directory, conocido también por sus siglas AD) es un servicio de directorios desarrolado por Microsoft que permite almacenar información como usuarios y dispositivos en una base de datos centralizada y jerárquica. AD brinda servicios como autenticación, políticas de acceso y administración de grupos.

Active Directory almacena información sobre objetos en la red y hace que esta información sea fácil de encontrar y usar para administradores y usuarios. Active Directory utiliza un almacén de datos estructurados como base para una organización lógica y jerárquica de la información del directorio.

La seguridad está integrada con Active Directory a través de la autenticación de inicio de sesión y el control de acceso a los objetos del directorio. Con un solo inicio de sesión en la red, los administradores pueden administrar los datos del directorio y la organización en toda su red, y los usuarios autorizados de la red pueden acceder a los recursos en cualquier lugar de la red. La administración basada en políticas facilita la gestión incluso de la red más compleja. [11]

Comparación entre LDAP y AD		
	LDAP	**AD**
Nombre Completo	Protocolo Ligero de Acceso a Directorios	Directorio Activo
Función	Protocolo	Proveedor de servicios de directorios
Standard	Código Abierto	Propietario
Sistemas Soportados	Multiplatforma: Windows, Linux, macos	Para aplicaciones y usuarios de Windows
Uso principal	Consultar y modificar elementos en proveedores de servicios de directorio	Proveer autenticación, políticas, administración de grupos y usuarios, y muchos otros servicios en forma de una base de datos de directorio

1.6. Keycloak

Keycloak es un software de código abierto que permite el *Single Sign-On* o Inicio de Sesión Único con *Identity Management* y *Access Management* para aplicaciones y servicios modernos. Esta herramienta facilita la protección de aplicaciones y servicios con poca o ninguna codificación. Un Proveedor de identidad (en inglés: *Identity Pro-*

vider, también conocido por sus siglas IdP), permite que una aplicación (a menudo llamada *Service Provider* o SP) delegue su autenticación. [12]

Este software está escrito en Java y es compatible de forma predeterminada con los protocolos de federación de identidad SAML v2 y OpenID Connect (OIDC) / OAuth2. Está bajo licencia de Apache y es mantenido por Red Hat. [12]

1.6.1. Características

Los usuarios se autentican en Keycloak en lugar de hacerlo en las aplicaciones. Esto significa que no es necesario que cada aplicación tenga un formulario de inicio de sesión, autentique a los usuarios o almacene sus datos. Una vez entren en Keycloak, los usuarios no tendrán que iniciar sesión en las demás aplicaciones conectadas al software.

Lo mismo sucede cuando un usuario cierra sesión. Keycloak ofrece cierre de sesión único, lo cual significa que los usuarios solo tienen que desconectarse en una de las aplicaciones para salir de su cuenta en el resto.

Otra prestación de Keycloak son las federaciones de usuarios, que facilitan la compatibilidad con LDAP y otros servidores de directorios activos. También admite la implementación de servicios propios para usuarios guardados en otros tipos de almacenamientos como en bases de datos relacionales.

Keycloak ofrece como herramienta una consola de administración de cuentas, a través de la cual los usuarios pueden administrar sus propias cuentas. Pueden actualizar su perfil, cambiar sus contraseñas y configurar la autenticación en dos pasos. También pueden administrar sus sesiones y visualizar el historial de su cuenta.

Otra característica es que es una herramienta extensible porque permite la eliminación, adición y modificación de las bases de datos de usuarios, los métodos de autenticación y los protocolos. Está basada en protocolos estándares y soportan OpenID Connect, OAuth 2.0 y SAML. [12]

Keycloak facilita añadir la autenticación y un servicio seguro a aplicaciones. Permite que los desarrolladores se centren en la funcionalidad empresarial al no tener que preocuparse por los aspectos de seguridad de la autenticación. También posibilita la unificación de los métodos de autenticación de distintas aplicaciones sin modificarlas.

Gluu es otra de las tecnologías que tienen prestaciones y ventajas similares a Keycloak. Es un servicio *Open Source* que soporta *SAML*, *OpenID Connect, SSO* y *OAuth 2.0*. Sin embargo Gluu es un sistema que requiere de 8 GB de RAM y 40 GB de espacio en disco, mientras que Keycloak solo necesita de 512 Mb de RAM y 1 GB de disco. Por ello Keycloak se ajusta más a los recursos que se tienen en la Universidad de La Habana.

1.6.2. JSON Web Token

JSON Web Token (abreviado JWT) es un estándar abierto basado en JSON propuesto por IETF (RFC 7519) para la creación de tokens de acceso que permiten la propagación de identidad y privilegios [13].

Esta tecnología define una forma compacta y autónoma para transmitir de forma segura información entre las partes como un objeto JSON. Esta información es verificada y confiable ya que se encuentra firmada digitalmente. Los JWT se pueden firmar usando un secreto (con el algoritmo HMAC) o un par de claves públicas / privadas usando RSA.

JSON Web Token es un método compacto y autónomo para transmitir información, se basa en una cadena de texto que tiene 3 partes (Header, payload, signature) codificadas en Base64, separadas por un punto que es estregado a los clientes de una API como llave de acceso. Keycloak utiliza JWT para transmitir una llave secreta de acceso a usuarios con privilegios. [14]

1.6.3. REST API

Se utiliza API REST como medio de prueba para ejecutar peticiones HTTP que posteriormente serán aseguradas con un Json Web Token emitido por Keycloak. La arquitectura REST enfoca a todo lo que lo conforma como un recurso. Los servicios web REST son livianos, altamente escalables y mantenibles. Se usan muy comúnmente para el intercambio de información, es el estándar más lógico, eficiente y generalizado en la creación de API para servicios de internet.

El siguiente esquema muestra como se realiza el flujo de información para autenticarse a través de Keycloak:

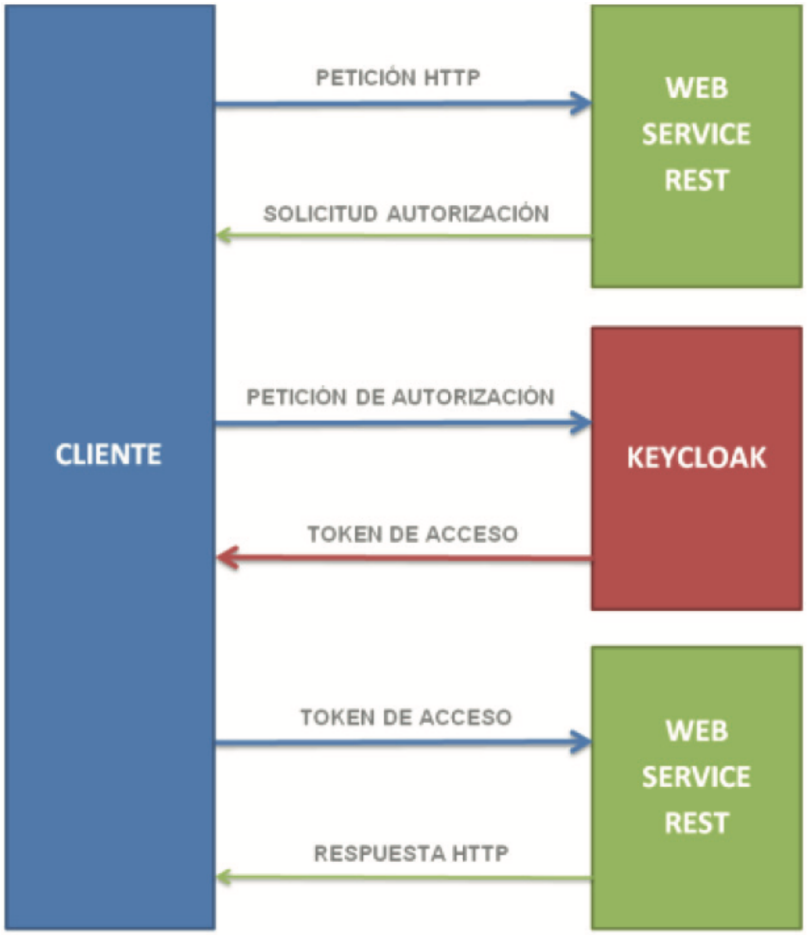

Figura 1.1

Este flujo asegurar que solo los usuarios a los que se les otorgó el acceso al servicio web puedan consumir dicha API. De esta forma se evita el acceso no autorizado a información que se transmite bajo un protocolo no seguro como lo es el protocolo HTTP. [14]

Capítulo 2

Propuesta

Capítulo 3

Detalles de Implementación y Experimentos

Conclusiones

Conclusiones

- El rendimiento del nuevo sistema es más eficiente ya que en cada iteración añade solo los nuevos usuarios en lugar de borrar y recrear cada noche todos sus usuarios. - El tiempo de activación de un usuario (el tiempo que transcurre entre que un usuario es dado de alta por recursos humanas y que el usuario tenga acceso a todos los servicios provistos por el Nodo Central como correo electrónico y acceso a internet) es más corto. - Sistema de autenticación más fiable

Recomendaciones

Recomendaciones

Bibliografía

[1] R. Kutera y W. Gryncewicz, «Single sign on as an effective way of managing user identity in distributed web systems. The ActGo-Gate project case study,» Informatyka Ekonomiczna, n.º 2 (40), 2016 (vid. págs. 1, 5).

[2] Microsoft. (2022). «What is single sign-on in Azure Active Directory?» Dirección: `https://learn.microsoft.com/es-es/azure/active-directory/manage-apps/what-is-single-sign-on` (vid. pág. 4).

[3] V. Radha y D. H. Reddy, «A survey on single sign-on techniques,» Procedia Technology, vol. 4, págs. 134-139, 2012 (vid. pág. 5).

[4] Google. (2022). «About SSO,» dirección: `https://support.google.com/` (vid. pág. 5).

[5] Auth0. (2022). «Single Sign-On,» dirección: `https://auth0.com/docs/authenticate/single-sign-on` (vid. pág. 5).

[6] OpenID. (2022). «Welcome to OpenID Connect,» dirección: `https://openid.net/connect/` (vid. pág. 5).

[7] C. Mainka, V. Mladenov, J. Schwenk y T. Wich, «Sok: Single sign-on security—an evaluation of openid connect,» en 2017 IEEE European Symposium on Security and P IEEE, 2017, págs. 251-266 (vid. pág. 5).

[8] R. Sánchez Guerrero, «Estudio y puesta en marcha de una infraestructura de gestión de identidad federada basada en SAML 2.0,» Tesis de mtría., 2009 (vid. pág. 6).

[9] N. Naik y P. Jenkins, «Securing digital identities in the cloud by selecting an apposite Federated Identity Management from SAML, OAuth and OpenID Connect,» en 2017 11th International Conference on Research Challenges in Information Science (IEEE, 2017, págs. 163-174 (vid. pág. 6).

[10] LDAP. (2022). «LDAP,» dirección: `https://ldap.com/` (vid. pág. 7).

[11] Microsoft. (2022). «Active Directory Domain Services Overview,» dirección: https://learn.microsoft.com/en-us/windows-server/identity/ad-ds/get-started/virtual-dc/active-directory-domain-services-overview (vid. pág. 8).

[12] Keycloak Documentation, https://www.keycloak.org/documentation (vid. pág. 9).

[13] J. Bradley, N. Sakimura y M. B. Jones, JSON Web Token (JWT), mayo de 2015. dirección: https://www.rfc-editor.org/rfc/rfc7519.txt (vid. pág. 10).

[14] C. Muyón y F. Montaluisa, «Métodos de seguridad de la información para proteger la comunicación y los datos de servicios web REST en peticiones HTTP utilizando JSON Web Token y Keycloak Red Hat Single Sign On,» Revista Ibérica de Sistemas e Tecnologias de Informação, n.º E29, págs. 198-213, 2020 (vid. págs. 10, 11).