



INSTITUTO POLITÉCNICO NACIONAL  
ESCUELA SUPERIOR DE CÓMPUTO

Tarea 1 - Resumen: Especificaciones URL y  
URI

Unidad de aprendizaje: Web Application Development

Grupo: 3CM6

*Alumnos(a):*

Ramos Diaz Enrique

*Profesor(a):*

Montes Casiano Hermes Francisco

15 de febrero 2019

# Índice

<b>1</b>	<b>Identificador Uniforme de Recursos (URI)</b>	<b>2</b>
1.1	Vista general	2
1.1.1	Uniform - Uniforme	2
1.1.2	Resource - Recurso	2
1.1.3	Identifier - Identificador	2
1.2	Identificación vs interacción	2
1.3	Sintaxis	3
1.4	Ejemplos	4
1.5	Caracteres reservados	4
1.6	Uso	4
1.6.1	Referencia de sufijo	4
<b>2</b>	<b>URI y URL (Localizador Uniforme de Recursos)</b>	<b>4</b>
2.1	Sintaxis general	5
2.2	Codificación de caracteres de URL	5
2.3	Esquemas específicos	5
<b>3</b>	<b>Bibliografía</b>	<b>5</b>

# 1. Identificador Uniforme de Recursos (URI)

Un *Identificador Uniforme de Recursos* (URI: Uniform Resource Identifier) proporciona un simple y extensible significado para identificar un recurso. Esta especificación de la sintaxis y semántica se deriva de los conceptos introducidos por la iniciativa global de información World Wide Web, cuyo uso data de 1990 y se describe en *Universal Resource Identifiers in WWW*.

## 1.1. Vista general

### 1.1.1. Uniform - Uniforme

La uniformidad permite usar diferentes tipos de identificadores de recursos al mismo tiempo, incluso cuando los mecanismos usados para acceder a estos son distintos, interpretar semánticas uniformes de convenciones sintácticas comunes en distintos tipos de identificadores de recursos, la introducción de nuevos tipos de identificadores de recursos sin interferir en la forma en que los existentes son utilizados, y reutilizar los identificadores en varios contextos distintos, permitiendo que nuevas aplicaciones o protocolos aprovechen un conjunto de identificadores de recursos existentes, de gran tamaño y ampliamente utilizados.

### 1.1.2. Resource - Recurso

El término *recurso* se utiliza en un sentido general para todo lo que pueda ser identificado por un URI. Ejemplos: un documento electrónico, un servicio, corporaciones, operadores de una ecuación matemática, tipos de relaciones en modelados de software, valores numéricos, etc.

### 1.1.3. Identifier - Identificador

Un identificador expresa la información requerida para distinguir lo que se está identificando de todas las otras cosas dentro de su alcance de identificación. Es decir, se refiere al propósito de distinguir un recurso de muchos otros disponibles, independientemente de cómo se logre este (puede ser por medio de su nombre, dirección, o contexto).

## 1.2. Identificación vs interacción

La propia URI solo proporciona identificación; el acceso al recurso no está garantizado ni implícito por la presencia de un URI. En su lugar, cualquier operación asociada con una referencia de URI se define por el elemento de protocolo, el atributo de forma de datos o el texto en lenguaje natural en el que aparece.

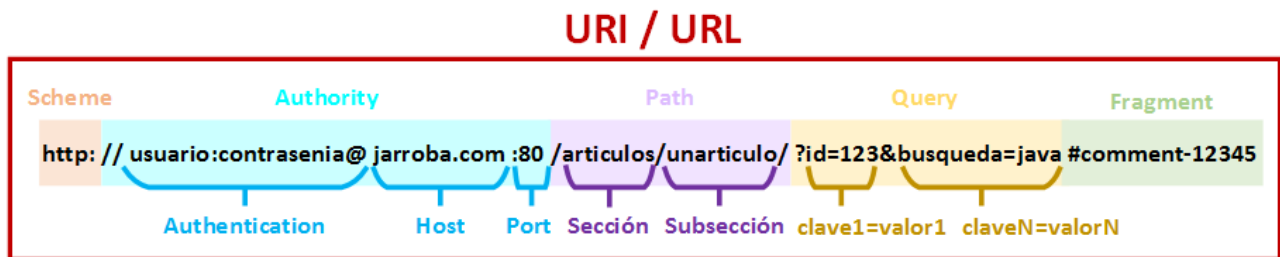
Dado un URI, un sistema puede intentar realizar una variedad de operaciones en el recurso. Estas operaciones están definidas por los protocolos que hacen uso de URI, no por su especificación.

La *resolución* de URI es el proceso de determinar un mecanismo de acceso y los parámetros apropiados necesarios para desreferenciar un URI.

### 1.3. Sintaxis

Un URI es un identificador que consiste en una secuencia de caracteres que consiste en una secuencia jerárquica de componentes referidos como el esquema, autoridad, ruta, consulta y fragmento.

La sintaxis de URI está organizada jerárquicamente, con componentes enumerados en orden decreciente de izquierda a derecha. Todo aquello después del delimitador de componente del esquema (":") se considera sin significado para el procesamiento del URI.



- **Esquema:** Cada URI comienza con un nombre de esquema que se refiere a una especificación para asignar identificadores dentro de ese esquema.
- **Autoridad:** Muchos esquemas de URI incluyen un elemento jerárquico para una autoridad de nombres, de modo que la gobernanza del espacio de nombres definido por el resto del URI se delega a esa autoridad (que a su vez, puede delegarlo).

Ejemplo: [ userinfo "@" ] host [ ":" port ]

- **UserInfo - Autenticación:** Puede consistir en un nombre de usuario e información específica del esquema sobre cómo obtener autorización para acceder al recurso. La información del usuario es seguida por un arroba ("@" ) que lo delimita del host.
- **Host:** Se identifica mediante un literal IP encapsulado entre corchetes, una dirección IPv4 en forma decimal con puntos o un nombre registrado.
- **Puerto:** Se designa mediante un número de puerto opcional en decimal que sigue al host y se delimita con un único carácter de dos puntos (":").
- **Ruta:** Contiene datos, generalmente organizados en forma jerárquica, que, junto con los datos en el componente de consulta, sirve para identificar un recurso dentro del alcance del esquema y la autoridad de denominación del URI. La ruta se termina con un "?", un "#", o cuando se llega al final del URI.
- **Consulta:** Contiene datos no jerárquicos que sirven para identificar un recurso dentro del alcance del esquema y la autoridad de denominación del URI (si corresponde). El componente de consulta inicia con un "?", y se termina con un "#", o cuando se llega al final del URI.
- **Fragmento:** Permite la identificación indirecta de un recurso secundario por referencia a un recurso primario e información de identificación adicional. Un componente fragmento se indica mediante la presencia de un "#", finaliza al final del URI.

## 1.4. Ejemplos

- ftp://ftp.is.co.za/rfc/rfc1808.txt
- http://www.ietf.org/rfc/rfc2396.txt
- ldap://[2001:db8::7]/c=GB?objectClass=one
- mailto:John.Doe@example.com
- news:comp.infosystems.www.servers.unix
- tel:+1-816-555-1212
- telnet://192.0.2.16:80/
- urn:oasis:names:specification:docbook:dtd:xml:4.1.2

## 1.5. Caracteres reservados

Los caracteres a continuación se denominan "reservados" porque pueden (o no) definirse como delimitadores por la sintaxis genérica, por cada sintaxis específica del esquema o por la sintaxis específica de la implementación del algoritmo de desreferenciación de un URI:

":", "/", "?", "#", "[", "]", "@", "!", "\$", "&", "'", "(", ")", "\*", "+", ",", ";", "=",

## 1.6. Uso

Cuando las aplicaciones hacen referencia a un URI, no siempre usan la forma de referencia completa definida por la regla de sintaxis "URI". Para ahorrar espacio y aprovechar la localidad jerárquica, muchos elementos de protocolo de Internet y formatos de tipo de medios permiten una abreviatura de un URI, mientras que otros restringen la sintaxis a una forma particular de URI.

### 1.6.1. Referencia de sufijo

La sintaxis URI está diseñada para una referencia inequívoca a los recursos y la extensibilidad a través del esquema URI. Sin embargo, como la identificación y el uso de URI se han convertido en algo común, los medios tradicionales han usado cada vez más un sufijo de URI como referencia, que consiste solo en las partes de autoridad y ruta de URI, tales como

www.w3.org/Addressing/

o simplemente un nombre DNS registrado por sí mismo.

## 2. URI y URL (Localizador Uniforme de Recursos)

Un URI puede clasificarse además como un localizador, un nombre o ambos. El término *Localizador Uniforme de Recursos* (URL: Uniform Resource Locator) se refiere al subconjunto de URI que, además de identificar un recurso, proporciona un medio para ubicar el recurso al describir su mecanismo de acceso primario (por ejemplo, su "ubicación" de red).

La especificación de las URL está diseñada para cumplir con los requisitos establecidos en *Functional Requirements for Internet Resource Locators*.

Las URL se utilizan para "localizar" los recursos, al proporcionar una identificación abstracta de la ubicación del recurso. Al ubicar un recurso, un sistema puede realizar una variedad de operaciones en el recurso, como podrían caracterizarse por palabras como "acceso", "actualización", "reemplazar", "encontrar atributos". En general, solo se debe especificar el método de "acceso" para cualquier esquema de URL.

## 2.1. Sintaxis general

En general, las URL se escriben de la siguiente manera: <esquema>:<parte-especifica-esquema>

Una URL contiene el nombre del esquema que esta siendo utilizado, seguido por dos puntos y luego una cadena de caracteres, cuya interpretación depende del esquema.

## 2.2. Codificación de caracteres de URL

Las URL se escriben solo con los caracteres gráficos imprimibles del conjunto de caracteres codificados US-ASCII. Los octetos 80-FF hexadecimal no se usan en US-ASCII, y los octetos 00-1F y 7F hexadecimal representan caracteres de control; estos deben estar codificados.

Los caracteres ";", "/", "?", ":", "@", "=", "&" pueden reservarse para un significado especial dentro de un esquema.

## 2.3. Esquemas específicos

El mapeo de algunos protocolos estándar y experimentales existentes se describe en la definición de sintaxis de BNF. Los esquemas cubiertos son:

- **ftp**: File Transfer Protocol - Protocolo de Transferencia de Archivos
- **http**: Hypertext Transfer Protocol - Protocolo de Transferencia de Hipertexto
- **mailto**: Direcciones de correo electrónicas
- **telnet**: Referencia a sesiones interactivas
- **news**: Noticias USENET
- **wais**: Servidores de información de área amplia
- **nntp**: Noticias USENET usando el acceso NNTP
- **file**: Nombres de archivos específicos del host

## 3. Bibliografía

[1] (2018) RFC 3986 Uniform Resource Identifier (URI): Generic Syntax. Accessed february 2019. [Online]. Available: <https://tools.ietf.org/html/rfc3986>

[2] (2018) RFC 1738 Uniform Resource Locators (URL). Accessed february 2019. [Online]. Available: <https://tools.ietf.org/html/rfc1738>