

Cómo los identificadores elegantes trabajan

Contenido

[Introducción](#)

[Números de placa de crédito](#)

U.S. [Números de la Seguridad Sociales](#)

[Números CUSIP](#)

[Números de la encaminamiento ABA](#)

Introducción

Este documento describe los identificadores elegantes, que son los modelos contenidos incorporados de la exploración que detectan los tipos determinados de datos. Para esta versión, el sistema implementará los identificadores elegantes para los números de placa de crédito, los números de Seguridad Social E.E.U.U., los números CUSIP, y los números de la encaminamiento ABA.

Internamente, un identificador elegante consiste en una expresión normal que haga juego las cadenas del candidato, junto con una función de la validación que marque la coincidencia del candidato de cierta manera. Por ejemplo, la función de la validación para un número de placa de crédito se asegura de que el dígito de control esté correcto.

Las expresiones normales para cada identificador elegante incluirán las anclas de la limitación de palabra ("b ") en los ambos extremos. (Esto evita que el sistema corresponda con un número de la Seguridad Social E.E.U.U., por ejemplo, en el medio de una cadena de dígitos más larga.) Para la simplicidad, éstos se omiten de las descripciones a continuación.

La implementación elegante de los identificadores debe tener cuidado sobre las coincidencias que solapan, porque una subcadena encontrada por la expresión normal puede no validar. Por ejemplo, un filtro está analizando para los números de placa de crédito contra la cadena 9999 que 4321 9999 9999 9995 1234 5678 9000 deben encontrar el número de placa de crédito válido 4321 9999 9999 9995, aunque una exploración simple de la expresión normal para los números posibles encontraría 9999 4321 9999 9999 y 9995 1234 5678 9000.

Números de placa de crédito

Un número de placa de crédito comienza con un tipo de placa de la Longitud variable, que indica si el número es una VISA, Mastercard, Amex, etc., y los extremos con un dígito de control. Diversos tipos de placa utilizan diversos números de dígitos en el número entero, pero el cálculo del dígito de control es lo mismo en cada caso.

Observe eso en el camino o los indicadores luminosos LED amarillo de la placa muestra gravedad menor JCB no se corresponden con. También, los números de la VISA 13-digit no

existen, y no serán correspondidos con en nuestra implementación.

los números de placa de crédito 16-digit harán juego una de las siguientes expresiones regulares:

```
[0-9]{4}-[0-9]{4}-[0-9]{4}-[0-9]{4}
[0-9]{4}\.[0-9]{4}\.[0-9]{4}\.[0-9]{4}
[0-9]{4} [0-9]{4} [0-9]{4} [0-9]{4}
[0-9]{16}
```

Con el prefijo siendo el "4", el "51"- "55", o el "6011".

Los números Amex 15-digit harán juego una de las siguientes expresiones regulares:

```
[0-9]{4}-[0-9]{6}-[0-9]{5}
[0-9]{4}\.[0-9]{6}\.[0-9]{5}
[0-9]{4} [0-9]{6} [0-9]{5}
[0-9]{15}
```

Con los prefijos permisibles siendo los "34" o los "37".

Los números del club de los comensales 14-digit harán juego una de las siguientes expresiones regulares:

```
[0-9]{4}-[0-9]{6}-[0-9]{4}
[0-9]{4}\.[0-9]{6}\.[0-9]{4}
[0-9]{4} [0-9]{6} [0-9]{4}
[0-9]{14}
```

Con los prefijos permisibles siendo los "300"- "305", los "36", o los "38".

Observe que las expresiones normales definen agrupar específico de los dígitos para una longitud de placa de crédito dada, y que si hay puntuación entre los dígitos, tiene que ser lo mismo en todas partes.

El dígito final en un número de placa de crédito es un dígito de control creado usando el algoritmo de Luhn. Trabajando del extremo derecho del número, doble cada segundo dígito. Entonces agregue para arriba los dígitos individuales de los números resultantes (ambos los que fueron doblados y los que no eran). Si el resultado es un múltiplo de 10, después el número es válido.

Por ejemplo, dado el número 1234 5678 9012 3456:

1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6

Doble: 2 2 6 4 10 6 14 8 18 0 2 2 6 4 10 6

Agregando 2 + 2 + 6 + 4 + 1 + 0... + 1 + 0 + 6 da 64, que no es un múltiplo de 10, así que el número es inválido.

Dado el número 1234 5678 9876 3333:

1 2 3 4 5 6 7 8 9 8 7 6 3 3 3 3

Doble: 2 2 6 4 10 6 14 8 18 8 14 6 6 3 6 3

Agregando $2 + 2 + 6 + 4 + 1 + 0... + 6 + 3$ da 80, que es un múltiplo de 10, así que el número es válido.

U.S. Números de la Seguridad Sociales

Los números de la Seguridad Sociales se dividen en un número de área de 3 dígitos, que se asigna geográficamente, número de grupo del dígito del a2 asignado en un orden particular dentro de un área, y un número de serie de 4 dígitos asignado secuencialmente.

Nuestra implementación utilizará las siguientes expresiones regulares:

```
[0-9]{3}-[0-9]{2}-[0-9]{4}
[0-9]{3}\.[0-9]{2}\.[0-9]{4}
[0-9]{3} [0-9]{2} [0-9]{4}
```

Aquí están algunos ejemplos de las expresiones arriba:

```
555-55-5555
555.55.5555
555 55 5555
```

La administración de Seguridad Social mantiene una lista del área/de los números de grupo se han asignado que: SSN publicado [3]. Pero puesto que este documento cambia periódicamente, no podemos confiar en él para la validación. La función de la validación marcará que ningunos de los 3 campos son todos los ceros, y que los primeros 3 dígitos son menos de 800. (Los usos de referencia anteriores 771 como el límite, pero el SSA tiene ya assigned number con los primeros 3 dígitos 771 y 772.)

(Los números que comienzan con 666 son no asignados, y los números en el rango 987-65-4320 a 987-65-4329 son reservados para hacer publicidad. También, 078-05-1120 es el SSN empleado mal; era el SSN real de una secretaria en una compañía de la cartera, que utilizó el número como un ejemplo.)

Números CUSIP

Los números CUSIP (comité sobre los procedimientos uniformes de la identificación de Seguridad) son 9 identificadores alfanuméricos que identifican las Seguridades norteamericanas de los diversos tipos. El número se divide en un número del emisor del 6-carácter, que identifica únicamente el emisor (e.g., compañía), el sufijo del carácter del a2 que identifica la Seguridad determinada; e.g., acción ordinaria, contra la acción preferida contra la opción contra el instrumento fijo de la renta.

El código elegante del identificador CUSIP utilizará las siguientes expresiones regulares:

```
[0-9]{3}[0-9a-zA-Z]{3} [0-9a-zA-Z]{2} [0-9]
[0-9]{3}[0-9a-zA-Z]{3}-[0-9a-zA-Z]{2}-[0-9]
[0-9]{3}[0-9a-zA-Z]{3}[0-9a-zA-Z]{2}[0-9]
```

La función de la validación es similar a la que está usada para los números de placa de crédito.

La única diferencia es que las cartas en el número CUSIP son convertidas a un valor numérico asignando A=10, B=11,..., Z=35.

Un ejemplo del sitio web de cusip.com utiliza el número 392690 cuarto de galón 3 CUSIP:

3 9 2 6 9 T3 0 Q

Cartas del convertido: 3 9 2 6 9 0 26 29 3

Doble: 3 18 2 12 9 0 26 58 3

Agregando $3 + 1 + 8 + 2 + 1 + 2 + \dots + 5 + 8 + 3$ da 50, que es un múltiplo de 10, así que el número original era válido.

Números de la encaminamiento ABA

Un número de la encaminamiento ABA (asociación bancaria americana) es un valor 9-digit. Los primeros 4 dígitos son Federal Reserve que rutea el símbolo, los 4 siguientes el identificador de la institución, y el último un dígito de control.

El código elegante del identificador del número de la encaminamiento ABA utilizará las siguientes expresiones regulares:

$[0-9]\{4\} [0-9]\{4\} [0-9]$
 $[0-9]\{4\}-[0-9]\{4\}-[0-9]$
 $[0-9]\{9\}$

La función de la validación implicada multiplicando cada dígito por 3, 7, 1,...; si la suma de los Productos es un múltiplo de 10, el número es válido.

Por ejemplo considere el número 123 456 789:

1 2 3 4 5 6 7 8 9

Multiplíquese por: 3 7 1 3 7 1 3 7 1

Producto: 3 14 3 12 35 6 21 56 9

Agregar $3 + 14 + 3 + 12 + 35 + 6 + 21 + 56 + 9$ da 159, que no es un múltiplo de 10, así que el número original era inválido.

Dado el número 322 271 627:

3 2 2 2 7 1 6 2 7

Multiplíquese por: 3 7 1 3 7 1 3 7 1

Producto: 9 14 2 6 49 1 18 14 7

Agregar $9 + 14 + 2 + 6 + 49 + 1 + 18 + 14 + 7$ da 120, que es un múltiplo de 10, así que el número original era válido.

(Aunque ciertos rangos de Federal Reserve que rutean los símbolos sean reservados, y por lo tanto no asignado, el algoritmo de la validación no marcará para saber si hay números

reservados, evitar tener que revisarlo si el ABA cambia su directiva.)