

스마트 식별자 작동 방식

목차

[소개](#)

[신용 카드 번호](#)

[미국 사회 보장 번호](#)

[CUSIP 번호](#)

[ABA 라우팅 번호](#)

소개

이 문서에서는 특정 유형의 데이터를 탐지하는 내장 콘텐츠 검사 패턴인 스마트 식별자에 대해 설명합니다. 이번 릴리스에서는 신용 카드 번호, 미국 사회 보장 번호, CUSIP 번호 및 ABA 라우팅 번호에 대한 스마트 식별자를 구현합니다.

내부적으로 스마트 식별자는 후보 문자열과 일치하는 정규식과 어떤 식으로든 후보 일치를 확인하는 검증 함수로 구성됩니다. 예를 들어 신용 카드 번호에 대한 검증 기능은 확인 번호가 정확한지 확인합니다.

각 스마트 식별자의 정규식은 양쪽 끝에 단어 경계 앵커('b')를 포함합니다. (이렇게 하면 시스템이 미국 사회 보장 번호(예: 긴 숫자 문자열 가운데)를 매칭하지 못하게 됩니다.) 간단한 설명을 위해 아래 설명에서 생략되었습니다.

스마트 식별자 구현은 정규식에서 찾은 하위 문자열이 검증되지 않을 수 있으므로 중복 일치에 대해 주의해야 합니다. 예를 들어, 필터는 9999 4321 9999 999 99995 1234 5678 9000에서 유효한 신용 카드 번호 4321 99999999999999999999999999999999 문자열에 대해 신용 카드 번호를 스캔하는 경우 99999999999999999999999999999999를 찾을 수 있습니다. 9 4321 999 9999 및 9995 1234 5678 9000

신용 카드 번호

신용 카드 번호는 가변 길이 카드 유형으로 시작하며, 이는 번호가 VISA, MasterCard, AMEX 등인지 여부를 나타내며, 수표 숫자로 끝납니다. 카드 유형마다 전체 숫자에 서로 다른 숫자를 사용하지만, 각 경우 확인 숫자 계산이 동일합니다.

enRoute 또는 JCB 카드가 일치하지 않습니다. 또한 13자리 VISA 번호는 존재하지 않으며 구현 시 일치하지 않습니다.

16자리 신용 카드 번호는 다음 정규식 중 하나와 일치합니다.

```
[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}
```

접두사는 "4", "51"- "55" 또는 "6011"입니다.

15자리 AMEX 번호는 다음 정규식 중 하나와 일치합니다.

[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}

허용되는 접두사는 "34" 또는 "37"입니다.

14자리 Diners Club 번호는 다음 정규식 중 하나와 일치합니다.

[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}

허용되는 접두사는 "300"- "305", "36" 또는 "38"입니다.

정규식은 지정된 신용 카드 길이에 대한 특정 숫자 그룹화를 정의하며, 숫자 사이에 구두점이 있는 경우 전체 범위에서 동일해야 합니다.

신용 카드 번호의 마지막 숫자는 Luhn 알고리즘을 사용하여 생성된 확인 숫자입니다.번호의 오른쪽 끝부터 두 번째 자리마다 두 번씩 작업을 수행합니다.그런 다음 결과 번호의 개별 숫자를 추가합니다(두 번 및 그렇지 않은 숫자 모두). 결과가 10의 배수인 경우 유효한 숫자입니다.

예를 들어 번호 1234 5678 9012 3456을 지정하면

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

이중:2 2 6 4 10 6 14 8 18 0 2 6 4 10 6

2 + 2 + 6 + 4 + 1 + 0 ... + 1 + 0 + 6을 추가하면 64가 되므로 10의 배수가 아니므로 번호가 유효하지 않습니다.

번호 1234 5678 9876 3333을 지정하면

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

이중:2 2 6 4 10 6 14 8 18 8 14 6 3 6 3

2 + 2 + 6 + 4 + 1 + 0 + 6 + 3을 추가하면 80이 제공되는데, 이는 10의 배수이므로 숫자가 유효합니다.

미국 사회 보장 번호

사회 보장 번호는 3자리 영역 번호로 구분되며, 이 번호는 지리적으로 할당되며, 한 영역 내의 특정 순서로 2자리 그룹 번호가 할당되고, 순차적으로 4자리 일련 번호가 할당됩니다.

구현에서는 다음 정규식을 사용합니다.

[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}

위의 식의 몇 가지 예는 다음과 같습니다.

555-55-5555
555.55.5555
555 55 5555

사회보장국은 할당된 지역/그룹 번호 목록을 유지 관리합니다.SSN이 [3]을(를) 발행했습니다.그러나 이 문서는 정기적으로 변경되므로 검증을 위해 이 문서에 의존할 수 없습니다. 검증 기능은 3개의 필드 중 모두 0이 아니며 처음 3개의 숫자가 800보다 작은지 확인합니다. 이전 참조에서는 771을 한도로 사용하지만 SSA는 이미 처음 3자리 숫자 771과 772의 숫자를 할당했습니다.

(666으로 시작하는 번호는 할당되지 않으며 987-65-4320~987-65-4329 범위의 번호는 광고용으로 예약되어 있습니다.또한 078-05-1120이 가장 잘못 사용되는 SSN입니다.전자지갑 회사의 비서관 SSN을 예로 들었다.

CUSIP 번호

CUSIP(Committee on Uniform Security Identification Procedures) 번호는 다양한 유형의 복미 증권을 식별하는 9개의 영숫자 식별자입니다.이 번호는 6자의 발급자 번호로 구분되며, 발급자(예: 회사)를 고유하게 식별하며, 특정 보안을 식별하는 2개의 문자 접미어입니다.예: 보통주 대 우선주 대 옵션 대 고정 수입 상품

CUSIP 스마트 식별자 코드는 다음과 같은 정규식을 사용합니다.

[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]

검증기능은 신용카드 번호에 사용하는 것과 유사합니다.유일한 차이점은 CUSIP 번호의 문자가 A=10, B=11, ..., Z=35를 할당하여 숫자 값으로 변환된다는 점입니다.

cusip.com 웹 사이트의 예는 CUSIP 번호 392690 QT 3을 사용합니다.

3 9 2 6 9 0 Q T 3

문자 변환:3 9 2 6 9 0 26 29 3

이중:3 18 2 12 9 0 26 58 3

3 + 1 + 8 + 2 + 1 + 2 + + 5 + 8 + 3을 추가하면 50이 되는데, 이는 10의 배수이므로 원래 번호가 유효했습니다.

ABA 라우팅 번호

ABA(American Banking Association) 라우팅 번호는 9자리 값입니다.처음 4자리는 연방 준비

(Federal Reserve) 라우팅 기호, 다음 4자리 기관 식별자 및 마지막 확인 숫자입니다.

ABA 라우팅 번호 스마트 식별자 코드는 다음과 같은 정규식을 사용합니다.

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

유효성 검사 기능은 각 숫자에 3, 7, 1, ...;제품의 합계가 10의 배수이면 유효한 값입니다.

예를 들어 숫자 123 456 789를 고려하십시오.

1 2 3 4 5 6 7 8 9

곱하기:3 7 1 3 7 1 3 7 1

제품:3 14 3 12 35 6 21 56 9

$3 + 14 + 3 + 12 + 35 + 6 + 21 + 56 + 9$ 를 추가하면 159가 되며 10의 배수가 아니므로 원래 번호가 유효하지 않습니다.

번호 322 271 627을 지정하면

3 2 2 2 7 1 6 2 7

곱하기:3 7 1 3 7 1 3 7 1

제품:9 14 2 6 49 1 18 14 7

$9 + 14 + 2 + 6 + 49 + 1 + 18 + 14 + 7$ 을 추가하면 120이 부여되는데, 이는 10의 배수이므로 원래 번호는 유효했습니다.

(특정 범위의 연방 예약 라우팅 기호는 예약되어 있으며, 따라서 지정되지 않더라도, ABA가 정책을 변경할 경우 이를 수정할 필요가 없도록 검증 알고리즘에서 예약된 번호를 확인하지 않습니다.)