Sistema de Identificación de Casos (Hash desde UUID)

Resumen

Este documento describe el sistema de identificación de casos para Botilito. El sistema proporciona IDs de caso legibles y fáciles de buscar, manteniendo los UUIDs como clave primaria en la base de datos.

Planteamiento del Problema

Situación Actual

- La base de datos usa UUIDs como claves primarias (ej: `a7b3c9e2-1234-5678-9abc-def012345678`)
- Los UUIDs no son amigables para referencia o comunicación verbal
- El `caseCodeGenerator.ts` existente genera IDs con secuencias aleatorias (001-999)
- Las secuencias aleatorias pueden colisionar cuando se crean múltiples casos simultáneamente

Objetivo

Los usuarios necesitan **encontrar y referenciar** casos usando un ID legible que:
- Sea fácil de comunicar verbalmente
- Contenga contexto significativo (tipo, fuente, fecha)
- Sea único garantizado (sin colisiones)
- Pueda usarse para búsqueda y filtrado

Solución: Sistema de ID Dual

Esquema de Base de Datos

```
┌─────────────────────────────────────────────────────────────┐
│ tabla cases                                                 │
├─────────────────────────────────────────────────────────────┤
│ id (UUID)            → Clave primaria, auto-generada        │
│ display_id (VARCHAR) → Legible, derivado del UUID, UNIQUE   │
│ ...otros campos...                                          │
└─────────────────────────────────────────────────────────────┘
```

Campo	Tipo	Propósito	Ejemplo
id	UUID	Clave primaria, relaciones, uso interno	a7b3c9e2-1234-567
display_id	VARCHAR(20)	Referencia para usuarios, búsqueda, compartir	T-WA-20241214-A7

Formato del Display ID

```
Formato: TIPO-VECTOR-FECHA-HASH

         T   -WA   -20241214-A7B
         ■    ■       ■      ■
         ■    ■       ■      ■■■ 3 caracteres del UUID (identificador único)
         ■    ■       ■■■ Fecha: AAAAMMDD
         ■    ■■■ Vector: Plataforma/fuente de transmisión
         ■■■ Tipo: Clasificación de contenido
```

Componentes

TIPO (Tipo de Contenido)

Código	Significado
T	Texto
I	Imagen
V	Video
A	Audio

Nota: Las URLs se clasifican como `T` *(texto)*

VECTOR (Fuente de Transmisión)

Código	Plataforma
WA	WhatsApp
FB	Facebook
TW	Twitter/X
IG	Instagram
TK	TikTok
YT	YouTube
TL	Telegram
WB	Web/Sitio
EM	Email
SM	SMS
OT	Otro

FECHA

- Formato: `AAAAMMDD`
- Representa la fecha de creación del caso

- Ejemplo: `20241214` = 14 de diciembre de 2024

HASH (Identificador Único)

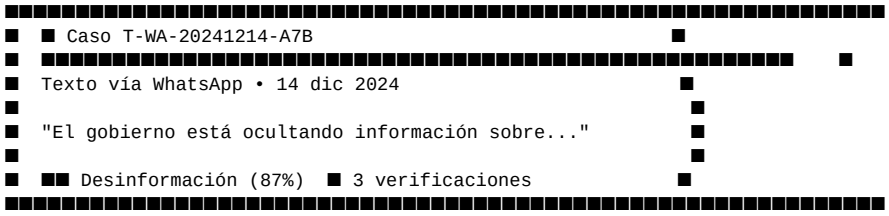
- 3 caracteres alfanuméricos derivados del UUID
- Usa los primeros 3 caracteres del UUID (sin guiones)
- Convertido a mayúsculas
- Garantiza unicidad (ya que el UUID es único)

Ejemplos

Registros en Base de Datos

id (UUID)	display_id	Descripción
a7b3c9e2-1234-5678-9abc-def012345678	T-WA-20241214-A7B	Texto de WhatsApp
k7m2p4x1-5678-9abc-def0-123456789abc	I-FB-20241214-K7M	Imagen de Facebook
p2x8n3q5-abcd-ef01-2345-6789abcdef01	T-TW-20241214-P2X	Texto de Twitter
9ab4c7e2-2468-1357-9bdf-ace024681357	V-TK-20241214-9AB	Video de TikTok
f3n6k9p2-9876-5432-1fed-cba987654321	A-TL-20241214-F3N	Audio de Telegram

Visualización para el Usuario



Casos de Uso

1. Referencia Verbal

```
Usuario A: "Oye, revisa el caso T-WA-20241214-A7B"
Usuario B: [busca T-WA-20241214-A7B, encuentra el caso al instante]
```

2. Búsqueda y Filtrado

```
Buscar por ID completo:  T-WA-20241214-A7B  → Coincidencia exacta
Filtrar por tipo:        T-*                → Todos los casos de texto
Filtrar por vector:      *-WA-*             → Todos los casos de WhatsApp
Filtrar por fecha:       *-20241214-*       → Todos los casos del 14 dic
```

```
Filtrar por hash:        *-A7B              → Encontrar por ID parcial
```

3. Compartir

```
URL: https://botilito.app/caso/T-WA-20241214-A7B
Copiar: "Caso T-WA-20241214-A7B"
```

4. Reportes y Estadísticas

```
"Hoy recibimos 47 casos de WhatsApp (T-WA-*)"
"Esta semana hubo 12 videos de TikTok (V-TK-*)"
```

Implementación Técnica

Algoritmo de Generación

```
Entrada:
  - uuid: "a7b3c9e2-1234-5678-9abc-def012345678"
  - contentType: "texto"
  - transmissionVector: "WhatsApp"
  - createdAt: 2024-12-14

Proceso:
  1. Mapear contentType → "T"
  2. Mapear transmissionVector → "WA"
  3. Formatear fecha → "20241214"
  4. Extraer hash: uuid.replace(/-/g, '').substring(0, 3).toUpperCase() → "A7B"
  5. Combinar: "T-WA-20241214-A7B"

Salida: "T-WA-20241214-A7B"
```

Algoritmo de Parsing

```
Entrada: "T-WA-20241214-A7B"

Proceso:
  1. Dividir por "-" → ["T", "WA", "20241214", "A7B"]
  2. Mapeo inverso "T" → "texto"
  3. Mapeo inverso "WA" → "WhatsApp"
  4. Parsear fecha "20241214" → 14 de diciembre, 2024
  5. Hash "A7B" usado para búsqueda

Salida:
  {
    contentType: "texto",
    transmissionVector: "WhatsApp",
    date: "2024-12-14",
    hash: "A7B"
  }
```

Consideraciones de Base de Datos

Definición de Columna

```
ALTER TABLE cases ADD COLUMN display_id VARCHAR(20) UNIQUE;
```

Índice para Búsqueda Rápida

```
CREATE INDEX idx_cases_display_id ON cases(display_id);
```

Momento de Generación

- `display_id` debe generarse al momento de inserción
- Requiere conocer: UUID, tipo de contenido, vector, fecha de creación
- Puede generarse en la capa de aplicación o como trigger de base de datos

Estrategia de Migración

Para Registros Existentes

1. Consultar todos los casos existentes con sus UUIDs
2. Generar `display_id` para cada uno basándose en:

- UUID (para el hash)
- Tipo de contenido (desde metadata)
- Vector de transmisión (desde metadata)
- Fecha de creación (desde `created_at`)

3. Actualizar registros con el `display_id` generado

Para Nuevos Registros

- Generar `display_id` al momento de inserción
- Incluir en el payload de inserción junto con otros campos

Comparación: Antes vs Después

Aspecto	Antes (Secuencia Aleatoria)	Después (Hash de UUID)
Formato	T-WA-20241214-001	T-WA-20241214-A7B
Fuente de secuencia	Aleatorio 001-999	Primeros 3 chars del UUID
Riesgo de colisión	■■ Posible	■ Imposible

Significado semántico	Ninguno (aleatorio)	Ninguno (hash)
Capacidad de búsqueda	■ Sí	■ Sí
Implementación	Simple	Simple
Tablas DB adicionales	Ninguna	Ninguna

Lo que Este Sistema NO Proporciona

- **Ordenamiento secuencial**: `A7B` no significa "caso #1" o "caso #2"
- **Indicación de volumen**: No se puede saber cuántos casos se crearon ese día
- **Ordenamiento cronológico por ID**: Se debe usar el timestamp `created_at` para ordenar

Estos fueron excluidos intencionalmente porque el objetivo es **búsqueda/referencia**, no conteo u ordenamiento.

Archivos a Modificar

1. `src/utils/caseCodeGenerator.ts`

- Actualizar `generateCaseCode()` para aceptar parámetro UUID
- Cambiar generación de secuencia de aleatoria a derivada del UUID

2. `src/types/botilito.ts`

- Asegurar que el campo `display_id` esté definido en los tipos relevantes

3. **Esquema de base de datos**

- Agregar columna `display_id` a la tabla de casos
- Agregar constraint único e índice

4. **API del Backend**

- Generar `display_id` al crear el caso
- Soportar búsqueda por `display_id`

5. **Componentes del Frontend**

- Mostrar `display_id` en lugar de UUID en la UI
- Soportar búsqueda por `display_id`

Apéndice: Flujo de Ejemplo Completo

```
1. Usuario envía texto de WhatsApp para análisis

2. Backend crea el caso:
   - Genera UUID: a7b3c9e2-1234-5678-9abc-def012345678
   - Detecta tipo de contenido: texto → T
   - Detecta vector: WhatsApp → WA
   - Obtiene fecha: 2024-12-14 → 20241214
   - Extrae hash: A7B
   - Crea display_id: T-WA-20241214-A7B

3. Registro en base de datos:
   {
     id: "a7b3c9e2-1234-5678-9abc-def012345678",
```

```
     display_id: "T-WA-20241214-A7B",
     content: "...",
     ...
   }
```

4. Usuario ve en la UI:
   "Caso T-WA-20241214-A7B"

5. Usuario comparte con colega:
   "Revisa el caso T-WA-20241214-A7B"

6. Colega busca:
   Entrada: "T-WA-20241214-A7B"
   Resultado: Caso encontrado ✓

Registro de Decisiones

Fecha	Decisión	Justificación
2024-12-10	Usar hash derivado de UUID en lugar de secuencia	Elimina riesgo de colisión
2024-12-10	Usar 3 caracteres para el hash	Suficientemente corto para referencia verbal, suficientemente u
2024-12-10	Mantener UUID como clave primaria	Integridad de base de datos, relaciones, práctica estándar
2024-12-10	Agregar display_id como campo separado	Separación de responsabilidades, indexable, buscable

Documento creado: 10 de diciembre, 2024
Estado: Planificación - Listo para implementación