

BASE DE CONHECIMENTO · DOCUMENTO PÚBLICO

GSM-7 ou Unicode no SMS: porque um acento divide a mensagem em mais partes

WhatsApp e APIs

CÓDIGO

RLX-MZ-KB-2157

VERSÃO

v2026.10.16

PAÍS · IDIOMA

Moçambique · Português

MozOut Moçambique

RLX Holdings, S.A. · Grupo RLX Holdings

NUIT 401 191 607 · NUEL 105 045 974 · CAE 62022

<https://www.meu.mozout.co.mz/knowledgebase/2157>

GSM-7 ou Unicode no SMS: porque um acento divide a mensagem em mais partes

WhatsApp e APIs Criado em 16 de outubro de 2026 3 min de leitura



A resposta curta (para programadores): um SMS leva um número fixo de bytes (140). Com o alfabeto **GSM-7** cabem 160 caracteres. Basta **um** carácter fora desse alfabeto (ã, õ, á, ê, um emoji, umas aspas curvas) para a mensagem inteira passar a **UCS-2** (Unicode), onde só cabem 70. Uma mensagem mais comprida é partida em pedaços que o telemóvel volta a juntar (SMS concatenado), e cada pedaço perde espaço para o cabeçalho que os liga: ficam 153 caracteres por parte em GSM-7 e 67 em UCS-2. Em português, palavras como «não», «código» ou «está» bastam para mudar a conta.

O que cada carácter faz à mensagem

O que escreve	O que acontece
Letras sem acento, algarismos, pontuação comum, e é, è, à, ù, ò, ñ, ü, Ç	GSM-7: 160 por mensagem. Estas letras acentuadas existem no alfabeto GSM.
ã, õ, á, í, ó, ú, â, ê, ô	Não existem no GSM-7: a mensagem toda passa a UCS-2, 70 por mensagem.
ç minúsculo	O GSM-7 tem o Ç. Alguns fornecedores trocam ç por Ç, outros passam a mensagem a Unicode: depende do fornecedor.
€ [] { } ^ ~ \	Estão na extensão do GSM-7: cada um conta como dois caracteres.
Emojis	UCS-2, e muitos ocupam duas posições, por isso gastam o dobro do espaço.
Aspas curvas, travessão longo e espaço não separável, colados de um documento	Fora do GSM-7: passam a mensagem a Unicode sem ninguém dar por isso.

➔ Como controlar o tamanho

1 Decida o alfabeto pela mensagem: nos avisos curtos e nos códigos, pode escrever sem os acentos que obrigam a Unicode; numa mensagem a clientes, aceite o Unicode e escreva mais curto. «nao» em vez de «não» passa num código, mas num aviso parece descuido.

2 Conte depois de preencher as variáveis, não com o modelo. Verifique no código se todos os caracteres estão no alfabeto GSM-7; se um não estiver, o limite é o do Unicode. Muitas APIs de fornecedores devolvem o número de partes na resposta: leia-o e registe-o.

3 Limpe o texto no código antes de enviar: aspas curvas para direitas, o travessão longo para hífen, o espaço não separável para espaço normal. Em PHP, por exemplo:

```
$t = str_replace(["\u{201C}", "\u{201D}", "\u{2018}", "\u{2019}", "\u{2014}", "\u{00A0}"], ['"', "'", '"', "'", "-", " "], $t);
```

4 No SMPP, a codificação é consigo. O campo **data_coding** diz como vai o texto: 0 é o alfabeto por omissão do centro de SMS (normalmente GSM-7) e 8 é UCS-2. Codifique o texto de acordo; um data_coding errado dá pontos de interrogação ou letras trocadas no telemóvel. As mensagens longas partem-se com um cabeçalho UDH (esm_class a indicá-lo) ou com o campo message_payload, conforme o fornecedor preferir.

5 Numa API HTTP, o fornecedor parte e codifica por si; alguns pedem que declare o texto como Unicode. Envie sempre em UTF-8 e leia na documentação dele o que faz aos caracteres que não conhece.

6 Teste num telemóvel a sério, com o texto real e um nome de cliente acentuado lá dentro.

! Um nome de cliente com acento num modelo contado em GSM-7 transforma uma mensagem em duas ou três.

Os nomes, as moradas e os nomes de produtos são as variáveis que mais vezes trazem Unicode sem aviso.

💡 Se o texto já sai com pontos de interrogação da base de dados, o problema está antes do SMS: **acentos que viram símbolos estranhos**. E o nome que aparece como remetente tem regras próprias: **SMS marketing e Sender ID**.

🔊 PRECISA DE AJUDA?

Os acentos já aparecem estragados no seu site ou na base de dados? Abra um pedido e diga-nos onde os vê.

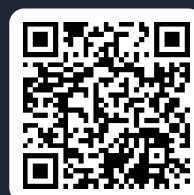
[📞 Abrir pedido de suporte →](#)**🔍 VEJA TAMBÉM**[SMS marketing: Sender ID, as regras daqui, e como escrever um que resulta →](#)[Acentos que viram símbolos estranhos no site: UTF-8 do ficheiro à base de dados →](#)[Conjuntos de caracteres: utf8mb4, ordenação e os pontos de interrogação no texto →](#)[SMPP ou API HTTP de SMS: o que é cada uma e de qual precisa a sua aplicação →](#)[Enviar códigos de uso único \(OTP\) por SMS com segurança: validade, tentativas e registos →](#)[WhatsApp, e-mail ou SMS: vantagens, desvantagens, riscos e para quem serve cada um →](#)**mozout****MozOut Moçambique**

RLX Holdings, S.A. · Grupo RLX Holdings

NUIT 401 191 607

NUEL 105 045 974

CAE 62022

SEDE Bairro 7 de Setembro
Vilankulo 1304 · CEP 0411-02, Inhambane, MoçambiqueCONTACTOS info@mozout.co.mz · +258 84 635 1111 · +258 86 635 1111 ·
www.mozout.co.mzHORÁRIO 2ª. à 6ª. Feira - 08 à 16 Horas
Sábado & Domingo - FechadoVersão online deste
documento

RLX-MZ-KB-2157 · v2026.10.16 · Gerado a 21/10/2026 17:54