

Base64-Codierung

Jacques Mock Schindler

19.03.2025

Die Base64-Codierung ist ein Verfahren zur Umwandlung von Binärdaten in eine Zeichenkette, die nur aus lesbaren ASCII-Zeichen besteht. Diese Codierung wird häufig verwendet, um binäre Daten über textbasierte Systeme zu übertragen, die möglicherweise nicht mit Binärdaten umgehen können (z.B. E-Mails).

Prinzip der Base64-Codierung

Base64 wandelt 3 Bytes (24 Bits) Binärdaten in 4 druckbare ASCII-Zeichen um:

1. Die Binärdaten werden in 6-Bit-Blöcke aufgeteilt ($2^6 = 64$ mögliche Werte)
2. Jeder 6-Bit-Block wird in ein druckbares ASCII-Zeichen umgewandelt
3. Bei unvollständigen Blocks am Ende werden Füllzeichen (=) hinzugefügt

Base64-Zeichensatz

Der Base64-Zeichensatz besteht aus 64 Zeichen:
- Die Großbuchstaben A-Z (26 Zeichen)
- Die Kleinbuchstaben a-z (26 Zeichen)
- Die Ziffern 0-9 (10 Zeichen)
- Zwei zusätzliche Zeichen, meist '+' und '=' (2 Zeichen)

Diese 64 Zeichen repräsentieren die Werte 0-63 und können mit 6 Bits dargestellt werden.

Codierungsprozess

1. Die Binärdaten werden in Gruppen von 3 Bytes (24 Bits) aufgeteilt
2. Diese 24 Bits werden in vier 6-Bit-Blöcke umgewandelt
3. Jeder 6-Bit-Wert wird als Index für den Base64-Zeichensatz verwendet

Wenn die Anzahl der zu codierenden Bytes nicht durch 3 teilbar ist:
- Bei einem übrig bleibenden Byte: Auffüllen mit vier Nullbits, Codierung ergibt zwei Zeichen und zwei '='-Zeichen
- Bei zwei übrig bleibenden Bytes: Auffüllen mit zwei Nullbits, Codierung ergibt drei Zeichen und ein '='-Zeichen

Beispiele:

Beispiel 1: Codierung von “KBW”

Buchstabe	K						B						W											
Dezimal	75						66						87											
Binär	0	1	0	0	1	0	1	1	0	1	0	0	0	1	0	0	1	0	1	1	1	1		
Dezimal	18						52						9						23					
Base64	S						0						J						X					

Figure 1: Base64 Beispiel

Beispiel 2: Codierung von “Hallo”

Betrachten wir das Wort “Hallo”:

1. ASCII-Werte: H=72, a=97, l=108, l=108, o=111
2. Binär: 01001000 01100001 01101100 01101100 01101111
3. In 6-Bit-Gruppen aufteilen: 010010 000110 000101 101100 011011 000110 1111
4. Da die letzte Gruppe nur 4 Bits hat, wird sie mit zwei Nullen aufgefüllt: 010010 000110 000101 101100 011011 000110 111100
5. Die 6-Bit-Werte sind: 18, 6, 5, 44, 27, 6, 60
6. Diese werden in Base64-Zeichen umgewandelt: S G F s b G 8 =

Das Ergebnis “SGFsbG8=” ist die Base64-Codierung von “Hallo”.

Anwendungen der Base64-Codierung

- E-Mail-Anhänge (MIME)
- Datenkodierung in URLs
- Einbettung von Bildern in HTML/CSS (Data-URLs)
- Übertragung binärer Daten in JSON
- Speicherung von Binärdaten in XML

Base64 vergrößert die Datenmenge um etwa 33% (da 3 Bytes in 4 Zeichen umgewandelt werden), bietet aber den Vorteil, dass die codierten Daten in jedem Textformat sicher übertragen werden können.