

A ChatGPT-nél a token egy olyan egység, amire a modell feldarabolja a szöveget, mielőtt feldolgozná vagy előállítaná. Mit jelent ez pontosan?

- A nyers szöveget (pl. amit te írsz be, vagy amit a modell ír ki) a rendszer nem **szavakban**, hanem kisebb darabokban tárolja.
- Ezek a darabok lehetnek:
 - egész szavak (pl. "kutya"),
 - szóelemek (pl. "számít", "ógép" két külön token),
 - néha csak betűk vagy szóközök is.

👉 Példa (angolul, mert ott jól látszik):

A mondat:

ChatGPT is smart.

Ez kb. így darabolódik tokenekre:

```
["Chat", "G", "PT", " is", " smart", "."]
```

Magyarban is hasonló a helyzet: a modell **nem mindig egész szavakat** tárol tokenként, hanem **darabokat**.

◆ Miért fontos?

- A **tokenek száma** határozza meg, mennyi szöveget tud egyszerre kezelni a modell. Például: ha egy modell 16 000 tokenet tud, az kb. **12–13 ezer szó** magyar szövegnek felel meg.
- A felhasználásnál is tokenekben mérik a költséget (pl. az OpenAI API-nál).

◆ Röviden:

A ChatGPT-ben a **token** a szöveg feldolgozásának alapegysége.

Nem pont szó, nem is betű, hanem valahol a kettő között: egy **szó- vagy szórészlet**.

Nézzük meg, hogyan darabolódik egy magyar mondat **ChatGPT tokenekre**.

(Az OpenAI modellek egy **Byte Pair Encoding – BPE** nevű algoritmust használnak, ami szavakat kisebb darabokra bont, ha kell.)

Példamondat:

Tokenizálás (hozzávetőlegesen így darabolja a modell):

- "A"
- " számít"
- "ógép"
- " gyorsan"
- " dolgozik"
- "."

🔑 Látható, hogy a „számítógép” kettészakadt:

- " számít" és
- "ógép"

Ez azért van, mert a modell a szövegek statisztikai gyakorisága alapján bontja fel a szavakat, nem pedig nyelvtan szerint.

Még egy példa:

Tokenek kb.:

- "Program"
- "ozni"
- " jó"
- " dolog"
- "."

Itt a „Programozni” két tokenből áll: "Program" + "ozni".

Tehát:

- Rövid, gyakran használt szavak sokszor **egy token** (jó, kutya).
- Hosszabb szavak több tokenre törnek (számítógép → 2 token).
- Ezért a magyar szöveg általában **több tokent fogyaszt**, mint ugyanaz angolul.

Értem, tehát arra vagy kíváncsi, hogy **milyen tényezők (pl. tokenek kezelése, beállítások, technikák)** segítenek abban, hogy a **ChatGPT jobb minőségű szöveget generáljon**. Ez nagyon jó kérdés, mert a "jobb szöveg" több dologon múlik:

🔑 1. Tokenek hossza és szerkezete

- A ChatGPT annál **pontosabban tud dolgozni**, minél **érthetőbb és tömörebb tokenláncot** kap.
- A rövid, egyértelmű mondatok **kevesebb félreértést okoznak**, így jobb minőségű a kimenet.
- Magyar nyelvben a hosszú szavak több tokent esznek → emiatt néha jobb **egyszerűbb fogalmazásokat** használni.

🔑 2. Prompt minősége (tehát amit beírsz)

A generált szöveg 70%-ban azon múlik, **hogyan kérdezel**.
Néhány tipp:

- **Légy konkrét:**
X „Írj valamit a kutyákról.”
✓ „Írj egy 5 mondatos, barátságos hangvételű szöveget a kutyák hűségéről, ami gyerekeknek is érthető.”

- **Adj példát:** Ha mutatsz mintát, a modell ahhoz fog igazodni.
 - **Stílusmegadás:** pl. „formális”, „vicces”, „tanulóknak”, „szakmai”.
-

3. Modellek és kapacitás

A ChatGPT több modellváltozatban érhető el:

- **GPT-3.5** → gyors, olcsóbb, de néha felszínesebb válaszok.
- **GPT-4 / GPT-5 (te most GPT-5-öt használsz nálam 😊)** → mélyebb gondolatmenet, jobb nyelvérzék, pontosabb stílus.
- A nagyobb modellek több **token** tudnak kezelni (pl. akár 100k tokenig), tehát **hosszabb szövegeket** is összefüggően értelmeznek.

4. Paraméterek (ha API-n használod)

- **temperature** → szabályozza a kreativitást
 - 0.2 körül: pontos, tényszerű, kevésbé változatos szöveg
 - 0.7–1.0: kreatívabb, változatosabb stílus
 - **max_tokens** → beállítja a válasz hosszát
 - **top_p** → a szavak valószínűségi szűrését finomítja
-

5. Iteráció (több körös finomítás)

Sokszor nem az első válasz a legjobb, hanem ha pontosítod:

- „Írd át rövidebbre.”
- „Fogalmazd meg gyerekeknek.”
- „Adj hozzá példákat.”

A modell ilyenkor **egyre jobb szöveget** fog generálni.

Összefoglalva:

A ChatGPT annál jobb szöveget ad, minél:

- tömörebben és egyértelműbben adod meg a bemenetet,
 - jól használod a tokentérfogatot,
 - tudatosan irányítod a stílust, hangnemet, hosszúságot,
 - és kihasználod a nagyobb modellek kapacitását.
 - Akkor vegyünk egy egyszerű témát: **„A Naprendszer”**, és nézzük meg, hogyan generál mást a ChatGPT különböző **prompt-beállítások** (stílus / célközönség) alapján.
-

-  **1. Gyerekbarát magyarázat**
- **Prompt:** „Magyarázd el a Naprendszert 8 éves gyerekeknek, mesés hangulatban.”
Válasz:
„A Naprendszer olyan, mint egy nagy család, ahol a Nap az anya, aki mindenkire vigyáz. Körülötte keringenek a bolygók, mint a gyerekek a játszótéren: a Merkúr a leggyorsabb, a Jupiter a legnagyobb, a Föld pedig a mi otthonunk. A Hold olyan, mint a Föld kísérő barátja, aki mindig mellette van.”

-  **2. Tudományos stílus**
- **Prompt:** „Írj egy rövid tudományos leírást a Naprendszerről középiskolásoknak.”
Válasz:
„A Naprendszer központi csillaga a Nap, amely gravitációjával maga körül tartja a bolygókat, holdakat, aszteroidákat és üstökösöket. A nyolc bolygó közül a belsők kőzetbolygók (Merkúr, Vénusz, Föld, Mars), a külsők pedig gáz- és jégóriások (Jupiter, Szaturnusz, Uránusz, Neptunusz). A rendszer kialakulása kb. 4,6 milliárd évvel ezelőttre tehető.”

-  **3. Vicces stílus**
- **Prompt:** „Írj vicces stílusban a Naprendszerről, mintha a bolygók beszélgetnének egymással.”
Válasz:
„A Nap azt mondja: »Hé, gyerekek, kicsit melegebb lesz ma!« Mire a Merkúr: »Na ne, én már így is grillezőlap vagyok!« A Jupiter büszkélkedik: »Én vagyok a legnagyobb, fogadjunk, hogy senki sem fér el mellettem.« A Föld mosolyog: »Nekem meg van pizza, meg wifi, úgyhogy mindenki nálam akar lakni.« A Szaturnusz pedig csak pörög a gyűrűjével, mintha divatbemutatón lenne.”

-  **4. Irodalmi, szépirodalmi stílus**
- **Prompt:** „Írj irodalmi hangvétellű leírást a Naprendszerről, mintha költemény lenne.”
Válasz:
„A Nap arany szívű úr, ki fényével eteti gyermekeit. A bolygók táncot járnak körülötte, mindegyik más dallamot dúdolva: a Föld a kék harmóniát, a Mars a vörös vágyat, a Jupiter mély mennydörgést, a Neptunusz halk, kék szonátát. Így szól a kozmosz örök kórusa, melynek mi is részei vagyunk.”

-  Ugyanaz a téma, **más prompt, más szöveg.**
Ezért van az, hogy a **promptolás művészete** (prompt engineering) az egyik kulcsa annak, hogy ChatGPT jobb szöveget adjon.

5 lépés a jobb szövegeneráláshoz

1. Legyél konkrét

- Ne kérj általánosságot, mert az általános választ ad.
X „Írj valamit a Naprendszerről.”

✓ „Írj egy 5 mondatos, barátságos stílusú szöveget a Naprendszeréről, amit 10 éves gyerekek is megértenek.”

2. Add meg a célt / közönséget

- Kinek szóljon a szöveg? (gyerek, laikus, szakember)
 - Milyen célra kell? (prezentáció, mese, tanulmány)
-

3. Határozd meg a formát

- Pl. felsorolás, esszé, vers, mese, dialógus.
 - ✓ „Írj rövid pontokban.”
 - ✓ „Fogalmazd meg meseként.”
 - ✓ „Írd le tudományos stílusban, lábjegyzetekkel.”
-

4. Adj példát vagy stílusmintát

- A ChatGPT remekül utánoz. Ha mutatsz egy stílust, abban folytatja.
Példa: „Így kezdődjön: »Egyszer volt, hol nem volt...«”
-

5. Finomíts iterációval

- Ha az első válasz nem tökéletes: pontosíts.
 - ✍ „Rövidítsd le felére.”
 - ✍ „Legyen viccesebb.”
 - ✍ „Használj egyszerűbb szavakat.”
-

⚡ Extra tipp:

Ha hosszabb munkát írsz (pl. esszé, cikk, tanulmány), bontsd részekre:

1. Vázlat → 2. Bevezető → 3. Részek kidolgozása → 4. Összefoglalás.
Így sokkal koherensebb lesz a végeredmény.

Univerzális Prompt Sablon

„Írj egy [hossz / formátum] szöveget a(z) [téma]-ról, amely [cél / közönség] számára készül.

Legyen a stílus [stílus / hangnem].

Tartalmazza a következőket: [kulcspontok / elvárások].”

 Példák kitöltve:

1. Gyerekeknek

„Írj egy **5 mondatos mesés hangulatú szöveget** a(z) **Nap energiájáról**, amely **10 éves gyerekeknek** szól.

Legyen a stílus **barátságos és játékos**.

Tartalmazza a következőket: **a Nap melege, fény, élet adása a Földön.**”

2. Tudományos szöveg

„Írj egy **rövid esszét** a(z) **klimaváltozás hatásairól**, amely **középiskolai diákoknak** szól.

Legyen a stílus **tudományos, de érthető**.

Tartalmazza a következőket: **hőmérséklet-emelkedés, olvadó jégsapkák, emberi felelősség.**”

3. Kreatív tartalom

„Írj egy **vicces párbeszédet** a(z) **Föld és a Mars között**, amely **szórakoztató blogbejegyzéshez** készül.

Legyen a stílus **humoros és könnyed**.

Tartalmazza a következőket: **éghajlat, lakhatóság, különbségek az emberek szemében.**”

 Ezt a sablont mindig az adott feladathoz igazíthatod.

Elég kicserélni a zárójeles részeket (**téma, közönség, stílus, formátum, kulcspontok**), és máris irányított, jó minőségű szöveget kapsz.