HTML

A **HTML** (angolul: *HyperText Markup Language*, „hiperszöveges jelölőnyelv”) egy [leíró nyelv](https://hu.wikipedia.org/wiki/Jel%C3%B6l%C5%91nyelv), melyet [weboldalak](https://hu.wikipedia.org/wiki/Weblap) készítéséhez fejlesztettek ki, és mára már internetes szabvánnyá vált a [W3C](https://hu.wikipedia.org/wiki/World_Wide_Web_Consortium) (World Wide Web Consortium) támogatásával. Az aktuális változata az 5, mely az [SGML](https://hu.wikipedia.org/wiki/Standard_Generalized_Markup_Language) általános [jelölőnyelv](https://hu.wikipedia.org/wiki/Jel%C3%B6l%C5%91nyelv) egy konkrét alkalmazása (azaz minden 5-ös HTML dokumentum egyben az SGML dokumentumszabványnak is meg kell hogy feleljen). Ezt tervek szerint lassan kiszorította volna az [XHTML](https://hu.wikipedia.org/wiki/HTML#XHTML), amely a szintén SGML alapú [XML](https://hu.wikipedia.org/wiki/XML) leíró nyelven alapul.

Használata

**HTML** általában szöveges állományokban található meg olyan [számítógépeken](https://hu.wikipedia.org/wiki/Sz%C3%A1m%C3%ADt%C3%B3g%C3%A9p), melyek az internethez kapcsolódnak. Ezek az állományok tartalmazzák azokat a szimbólumokat, amelyek a megjelenítő programnak leírják, hogyan is kell megjeleníteni illetve feldolgozni az adott állomány tartalmát. Megjelenítő program lehet egy [webböngésző](https://hu.wikipedia.org/wiki/Webb%C3%B6ng%C3%A9sz%C5%91) (angolul: web browser), aural böngésző (olyan, amelyik a felhasználónak felolvassa a megjelenítendő szöveget), [braille](https://hu.wikipedia.org/wiki/Braille-%C3%ADr%C3%A1s" \o "Braille-írás) olvasó, amely konvertálja a szöveget braille „formátumba”, [levelező program](https://hu.wikipedia.org/w/index.php?title=Levelez%C5%91_program&action=edit&redlink=1) (mint például: [Mozilla Thunderbird](https://hu.wikipedia.org/wiki/Mozilla_Thunderbird), [Microsoft Outlook](https://hu.wikipedia.org/wiki/Microsoft_Outlook), [Eudora](https://hu.wikipedia.org/w/index.php?title=Eudora&action=edit&redlink=1" \o "Eudora (a lap nem létezik)) stb.), valamint egyéb eszközök, például [mobiltelefon](https://hu.wikipedia.org/wiki/Mobiltelefon).

Szimbólumok, tartalom

Négyfajta szimbólum ([leíró elem](https://hu.wikipedia.org/wiki/HTML-elem)) található meg a **HTML**-ben:

* *strukturális* elemek, amelyek leírják az adott szöveg "célját" például *h1* mint első szintű címsor (alcím).
* *prezentációs* szimbólumok, amelyek leírják, hogy az adott szöveg hogyan nézzen ki: például *b* **vastag** kinézetet eredményez. (Ez a forma azonban ma már *elavultnak* számít, helyette a [CSS](https://hu.wikipedia.org/wiki/CSS) használata javasolt, ugyanis a legújabb irányelv szerint szét kell választani a tartalmat (amit a HTML kódol) és a formát (amit [CSS](https://hu.wikipedia.org/wiki/CSS)-ben szokás kódolni), és ezáltal mindenki számára elérhetővé tehetőek a weblapok: felolvasó gépeknek, amit a gyengén látók használnak, szöveges böngészőknek, és nyomtatóbarát verziót is könnyebb készíteni.)
* [*hiperszöveg*](https://hu.wikipedia.org/wiki/Hipersz%C3%B6veg)*(hypertext)* elemek, amelyek segítségével kapcsolat létesíthető a dokumentum egyes elemei és más dokumentumok között (például az a tag)
* *eszköz* elemek, amelyek segítségével gombok, listák, beviteli mezők hozhatók létre.

Felépítés

Egy HTML állomány három fő részre bontható:

1. A [Dokumentum Típus Definíció](https://hu.wikipedia.org/wiki/DTD) az állomány legelején, pl:  
   *doctype*
2. a HTML fejléc *head*, ami technikai és dokumentációs adatokat tartalmaz, amelyeket az internet böngésző nem jelenít meg, tehát átlag felhasználó ezeket nem látja és
3. a HTML törzs *body*, amely a megjelenítendő információkat tartalmazza.

**Karakterkódolás**

Egy HTML dokumentum alapértelmezésként [ISO-8859-1](https://hu.wikipedia.org/w/index.php?title=ISO-8859-1&action=edit&redlink=1), azaz nyugat-európai kódolást használ.[[1]](https://hu.wikipedia.org/wiki/HTML#cite_note-1) Gyakran előforduló hiba szokott lenni, hogy nincs beállítva a charset paraméter a fejléc content attribútumában, annak ellenére, hogy a dokumentum nem nyugat-európai kódolású szöveget tartalmaz. Magyar nyelvű oldalak esetén ilyenkor gyakran az „*ő*” és „*ű*” betűk helyett „*õ*” és „*û*” karaktereket látunk (ekkor valószínűleg "[**ISO-8859-2**](https://hu.wikipedia.org/wiki/ISO-8859-2)" vagy "**windows-1250**" volna a helyes charset érték), de az is lehet, hogy minden ékezetes betű teljesen olvashatatlanná válik, [betűszemetet](https://hu.wikipedia.org/wiki/Bet%C5%B1szem%C3%A9t) eredményezve (ha például a megjelenítendő szöveg [UTF-8](https://hu.wikipedia.org/wiki/UTF-8) kódolású). A charset paraméter értékeként több száz kódolás és érvényes alternatív név (alias) megadható.[[2]](https://hu.wikipedia.org/wiki/HTML#cite_note-2)

**Numerikus karakter referenciák**

Akármilyen kódolást használ is egy HTML dokumentum, bármilyen [Unicode](https://hu.wikipedia.org/wiki/Unicode) karakter megjeleníthető a szövegben annak [UCS](https://hu.wikipedia.org/wiki/UCS) kódjával „&#” és „;” jelek között. Például a magyar kis „ő” betű UCS kódja a [decimális](https://hu.wikipedia.org/wiki/Decim%C3%A1lis) 337 (vagy [hexadecimális](https://hu.wikipedia.org/wiki/Hexadecim%C3%A1lis) 151), ezért függetlenül a használt kódolástól, a HTML dokumentumban elhelyezett &#337; vagy &#x151; mindig az „ő” betűt fogja megjeleníteni.

**Karakter egyedhivatkozások**

Az [**SGML**](https://hu.wikipedia.org/wiki/SGML) számos [**Unicode**](https://hu.wikipedia.org/wiki/Unicode) karakterhez definiál egyedi megnevezéseket,[[3]](https://hu.wikipedia.org/wiki/HTML#cite_note-3)[[4]](https://hu.wikipedia.org/wiki/HTML#cite_note-4) azonban a HTML a lehetséges SGML egyedeknek csak egy részét használja[[5]](https://hu.wikipedia.org/wiki/HTML#cite_note-5) (a használt egyedeket a dokumentum [DTD](https://hu.wikipedia.org/wiki/DTD)-jében kell megadni, és a HTML 4.01-es verziójáig ez nem tartalmazza a közép-európai karaktereket). Emiatt a magyar ékezetes karakterek közül is csak azokra lehetséges SGML egyedhivatkozást megadni, amelyek a nyugat-európai kódkészletben is megtalálhatók. Vagyis hiába definiálja az SGML például „ő” betűhöz az „odblac” nevet, hivatalosan azt egy böngészőnek nem kell tudnia megjeleníteni egy 4.01-es HTML dokumentum esetén.

XHTML

Az XHTML a HTML megfogalmazása [XML](https://hu.wikipedia.org/wiki/XML)-ben. Gyakorlatilag nincs jelentős eltérés a két nyelv között, csak a formai követelmények lettek szigorúbbak:

* Mindent kisbetűvel kell írni! (kivéve a DOCTYPE elemet)
* Minden elemet le kell zárni! Az üres elemeket önmagukban egy szóközzel és egy / jellel.
* Az elemeket csak egymásba ágyazva lehet használni!
* A jellemzőket idézőjelek közé írjuk!
* A jellemzőknek legyen értéke!

HTML5

A **HTML5** a következő, jelentősen átdolgozott változata a HTML-nek (Hypertext Markup Language), a web fő jelölőnyelvének. Egyik fő tervezési célja, hogy a webes alkalmazásokhoz ne legyen szükség pluginek (pl. Adobe Flash, Microsoft Silverlight, Sun JavaFX) telepítésére. A specifikáció a HTML4 és az XHTML1 új verzióját jelenti, a hozzájuk tartozó DOM2 HTML API-val együtt. A HTML5 specifikációban leírt formátumba történő migráció HTML4-ről, vagy XHTML1-ről a legtöbb esetben egyszerű, mivel a visszamenőleges kompatibilitás biztosított. A specifikáció a közeljövőben támogatni fogja a Web Forms 2.0 specifikációt is. HTML5-ben lehetőség van audio és videofájlok beillesztésére, erre a célra külön tagek vannak. Továbbá létezik egy úgynevezett canvas tag, mely egy vászon. Erre [Javascript](https://hu.wikipedia.org/wiki/Javascript" \o "Javascript) segítségével lehet rajzolni, vagy képet beilleszteni. Fontos megemlíteni, hogy a HTML5-ből kivették az elrendezés megadó tulajdonságokat. Ezeket [CSS](https://hu.wikipedia.org/wiki/CSS)-ben kell megadni.