



### **Mi ez pontosan?**

A helyesírási tanácsadó portál egy olyan interneten keresztül elérhető szoftvereszköz, amely segít *a mindenkori helyesírási szabályzat* szerinti helyesen írott alak megtalálásában. A megoldás mellett magyarázatot is ad, ezzel is igyekszik elősegíteni a tudatos helyesírási készségek elsajátítását.

Az MTA Nyelvtudományi Intézetben többéves munkával (kb. 8 emberévnyi ráfordítással) készült rendszert a nagyközönség ingyenesen veheti használatba.

### **Miért készült?**

A nyelvi tanácsadás az Intézet fontos társadalmi küldetése az alapítása óta, ezt tehát immár több mint fél évszázada végez. Az utóbbi években felismertük, hogy a meglévő eszközökkel és kapacitással munkatársaink nem győzik e feladat ellátását. A telefonos szolgálat időben korlátozott, a hívás nem díjmentes, és már csak ezért sem érhető el mindenki számára. Az e-mailben feltett kérdésekre a legjobb igyekezet mellett sem adhatunk azonnali választ.

A változást az is indokoltta, hogy időközben megváltozott a környezet, amelyben a tanácsadó szolgálat működik. Az internet alapvető fejlődést eredményezett: elterjedésével sokkal több emberhez juttatható el a tanácsadó szolgáltatás. Az okostelefonok rohamos terjedése egészen új távlatokat nyit e téren is.

Mindezek megértették a felismerést, hogy korszerű technológián alapuló webszolgáltatással egészítsük ki a nyelvi tanácsadó szolgáltatást. Az online szolgáltatás előnye, hogy időben és térben korlátok nélkül elérhető, azaz folyamatosan és egyszerre rengeteg ember számára tudunk tanácsot adni, határon belül és külföldre egyaránt. Fontos különbség, hogy azonnali válasz nyújtható, és mindez nem kerül külön pénzbe.

A nyelvi tanácsadó szolgáltatás első lépcsőjeként a kérdések zömét kitevő helyesírás kérdéskörében adunk felvilágosítást. A most indult portál a későbbiekben kiterjed majd a nyelvi tanácsadás egyéb területeire is.

### **Miért nehéz?**

A helyesírás az embereknek sem könnyű, sokan tartanak tőle, és kevesen művelik helyesen. A szabályok értelmezése és alkalmazása mindenkit nehéz feladat elé állít.

Az automatizálás nagy szakmai kihívás. A helyes alak megadása ugyanis nem csak az alakon múlik. A rövid *i*-vel írt *igér* alakról önmagában véve el tudjuk dönteni, hogy helytelen. Gyakran azonban mindkét alak helyes, de csak a megfelelő környezetben, amelyeket nem egyszerű megkülönböztetni vagy megjegyezni (pl. *egyelőre* – *egyenlőre*, *helység* – *helyiség*). A számítógép számára azonban a legnehezebb feladat az, hogy értelmezni tudjuk a kifejezést, illetve a szándékolt jelentést egy adott környezetben (l. *orosz tanár* – *orosztanár*). A szöveg ilyen fokú megértése már a mesterséges intelligencia határát súroló feladat, amelyet a nyelvtechnológia jelenlegi fejlettségi szintjén nem tudunk maradéktalanul megoldani.



A nehézség másik forrása az, hogy a helyesírás veszélyes terep. A focihoz hasonlóan képes komoly indulatokat kiváltani. Komoly felelősséget jelent tehát kiállni a nagyközönség elé az MTA nevében egy számítógépes algoritmusokra, nyelvtechnológiára építő szolgáltatással.

### **Kiknek készült?**

Mottónk: helyesírás mindenkinek!

Felfogásunk szerint a helyesírás társadalmi követelmény, a társadalmi elfogadottság egyik mércéje, amelynek hiánya megbélyegzéshez és frusztrációhoz vezethet. A társadalmi nyilvánosság az internet jóvoltából lényegesen kitágult: tömegesen jutnak szóhoz olyanok, akiknek korábban nem volt rá lehetőségük. A közösségi média, a blogok, a hozzászólások elsősorban írásos közeget jelentenek. Kijelenthetjük tehát, hogy a helyesírás többé már nem csak kulturális elit réteg ügye. A helyesírás, a nyelvhelyesség támogatása egyben a nyelvi esélyegyenlőség elősegítését is jelenti. Ehhez a fontos társadalmi küldetéshez is hozzá szeretnénk járulni ezzel az eszközzel.

A fenti mottó alatt nem csak azt értjük, hogy ezzel a portállal megszüntetjük a tanácsadó szolgáltatásunk idő- és térbeli korlátait, és mindenhova elvisszük a helyesírást, ahol csak van internet. Ezen felül célunk az, hogy mindenki számára könnyen hozzáférhető, könnyen használható és főleg *könnyen érthető* eszközt készítsünk. Olyan eszközt, amely a nyelvileg kevésbé tudatos használók számára is hasznos. Ez nemcsak azt jelenti, hogy a rendszer barátságos, könnyen kiismerhető és használható legyen, hanem azt is hogy az ő nyelvükön szóljon, abban az értelemben, hogy ne csak a köznyelvi sztenderd helyesírásával foglalkozzon. A helyesírási szabályzat szándékosan nem foglalkozik nem sztenderd nyelvhasználattal mint például a laza beszélt nyelvi alakokkal vagy a tájnyelvi változatokkal.

Úgy véljük, hogy a digitális eszközök révén tömegessé vált kommunikáció kikényszeríti annak a távlati célnak a megvalósítását, hogy a rendszer használói a saját természetes nyelvhasználatuk szerint leírt alakokat adjanak meg, és kapjanak választ a köznyelvi norma szerint helyesen írt alakban.

### **Hogyan készült?**

A rendszerbe igyekeztük beépíteni azt a több évtizedes tudást, amely a nyelvi tanácsadó szolgálat működése alatt az Intézetben felhalmozódott. Feldolgoztuk a szolgálat naplóit a leggyakrabban kért kifejezésekről, ezeket az Archívum rovatban a portálon is közzétettük, valamint mostantól folyamatosan közöljük is.

Ezen felül a rendszer nemzetközi összehasonlításban is értékes jellemzője a korszerű nyelvtechnológia alkalmazása: a portál működéséhez több százezres tételt tartalmazó adatbázisokat, az egymilliárd szavas Magyar nemzeti szövegtár új kiadását, külön erre a célra készített, szemantikai jegyekkel ellátott szótárakat készítettünk, illetve használtunk fel. Mindezekre a nyelvi erőforrásokra támaszkodnak a háttérben futó programok, amelyek elemzik a kért kifejezést, és ennek alapján adják meg a választ.

### **Mitől intelligens?**

A rendszer legfontosabb újdonsága és a magyar nyelvi megoldásokat tekintve páratlan sajátossága, hogy nem egyszerűen szótári egybevetést végez. A megadott szóalakot, kifejezést megpróbálja



értelmezni, például. felismeri az anyagneveket, színneveket, a többértelmű alakokat (*megint – meg+int*). Az elemzés megtalálja a megfelelő szabályt, majd ennek alapján megadja a helyes alakot. Nemcsak javít, hanem magyaráz is, megadja azt a szabályt, amelynek alapján az alakot a javasolt módon kell írni, vagy elmagyarázza a két alakváltozat közötti különbséget. Büszkeségünk, hogy a rendszer olyan elemző szabályrendszert használ, amely elvileg tetszőleges számú hasonló kifejezést képes felismerni. Listába nem foglalható, végtelen elemszámú kifejezést képes tehát kezelni. Ez önmagában biztosítja előnyét a véges méretű szótárakat használó megoldásokkal szemben.

### **Mindent tud?**

Ma még sajnos nem... Ez az 1.0-s induló változat, amely hét kérdéskörben ad választ:

- különírás – egybeírás,
- helyes-e így?,
- névkereső,
- elválasztás,
- számok betűzése,
- dátumok betűzése,
- ábécébe rendezés.

A technológia ígéretes ugyan, de mint jeleztük, vannak korlátai. Elsősorban a szöveg megértésében, a jelentés megragadásában mutatkoznak a korlátok.

Ugyanakkor már csak azért sem beszélhetünk végleges változatról, mert maga a terület, a nyelvhasználat állandó változásban van, a szabályzat is – kellő mértéktartással – követi a változásokat. Ennél fogva szükséges a nyelvhasználat folyamatos vizsgálata, a nyelvtan és az adatbázisok karbantartása.

Mindent ugyan nem tudunk, de örömmel jelentjük, hogy **a rendszerünk 95 %-os eredményt ért el a [www.magyarhelyesírás.hu](http://www.magyarhelyesiras.hu) portálon található helyesírási teszteken.**

### **Lesz jobb is?**

Feltétlenül!

Terveink szerint a jövőendő változatok

- közérthetőbbek lesznek (esetleg külön készítünk egy gyors, kezdőknek és egy szakértőknek, vajt fülüknek szóló változatot);
- rugalmasabban működnek: kezelik majd a nem standard (beszélt nyelvi és nyelvjárási) változatokat;
- interaktívak lesznek: amit ma még gépi eljárással nem tudunk megérteni, a rendszer dialógus formájában megpróbálja a használatól megtudni.

### **Hogyan használhatom?**

Számítógépen, lapon és táblagépen a [helyesírás.mta.hu](http://helyesiras.mta.hu) vagy [helyesiras.mta.hu](http://helyesiras.mta.hu) címeiken. (Az [mta.helyesiras.hu](http://mta.helyesiras.hu), illetve a [helyesiras.hu](http://helyesiras.hu) címek, valamint mindezek [www](http://www) előtagú párja is működik.) És ugyanezek a címeiken használható a rendszer **okostelefonokon** is.