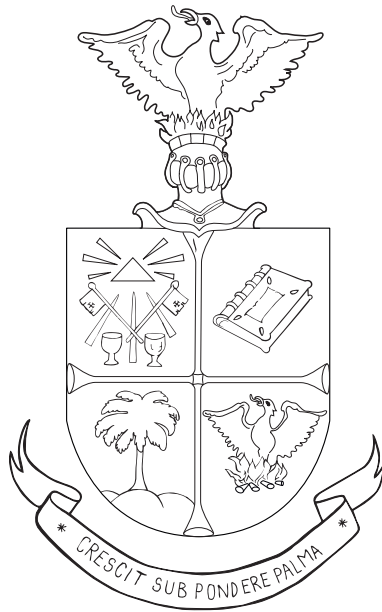


Glossa Iuridica

XI. évfolyam, 1-2. szám



Budapest, 2024

Károli Gáspár Református Egyetem Állam- és Jogtudományi Kar

SZERKESZTŐBIZOTTSÁGI ELNÖK:

Miskolczi Bodnár Péter

A SZERKESZTŐBIZOTTSÁG TAGJAI:

Antalóczy Péter, Birher Nándor, Böcskei Elvira, Bruhács János,
Csáki-Hatalovics Gyula, Cservák Csaba, Czine Ágnes, Dobrocsi Szilvia,
Domokos Andrea, Homicskó Árpád Olivér, Jakab Éva, Köbel Szilvia,
Kukorelli István, Kun Attila, Liviu Radu, Miskolczi Bodnár Péter, Móré Sándor,
Osztovits András, Pulay Gyula, Rixer Ádám, Erik Stenpien, Stipta István,
Szabó Imre, Szalma József, Szuchy Róbert, Tamás András, Tóth András,
Tóth J. Zoltán, Törő Csaba Attila, Trócsányi László, Udvary Sándor

FŐSZERKESZTŐ:

Rixer Ádám

FŐSZERKESZTŐ-HELYETTESEK:

Homicskó Árpád Olivér, Szuchy Róbert, Udvary Sándor

SZERKESZTŐK:

Boóc Ádám, Cservák Csaba, Domokos Andrea, Stipta István

ISSN 2064-6887

Kiadja a Károli Gáspár Református Egyetem Állam- és Jogtudományi Kara

Felelős kiadó: Prof. Dr. Tóth J. Zoltán, dékán

A kiadvány nyomdai munkálatait előkészítette:

Patrocinium Kiadó

A MAGYAR TERMINOLÓGIASZTRATÉGIA MEGVALÓSÍTHATÓSÁGA

A SZOFTVERES KERETRENDSZER KIVÁLASZTÁSÁNAK SZEMPONTJAI
A MAGYAR NEMZETI TERMINOLÓGIAI ADATBÁZIS TERVEZÉSE SORÁN¹

FÓRIS ÁGOTA
egyetemi tanár (KRE BTK)

SOMOGYI ZOLTÁN
nyelvi szakértő (Oracle Global Services Hungary)

Absztrakt

Az utóbbi időszakban fontos kérdéssé vált egyrészt az általános magyar terminológiastratégia megvalósíthatósága, másrészt egy jogi-közigazgatási terminológiastratégia elindítása. A terminológiastratégia megvalósítása során az egyik legfontosabb a terminológiai adatok összegyűjtése, ezek terminológiai (nyelvészeti és szakmai) szempontú feldolgozása, nyilvános közzétételük terminológiai adatbázisokban.

Az általános magyar terminológiastratégia egyik kitűzött feladata a Magyar Nemzeti Terminológiai Adatbázis (MaNTA) megtervezése és felépítése. E tanulmányban felmérjük az adatbázis kezelésére alkalmas, nyilvánosan elérhető keretrendszereket, azok funkcionalitását és lehetőségeit. A tanulmányban először a szoftveres keretrendszer kiválasztásának szempontjait vesszük számba, majd két adatbázis-kezelő szoftver, az ingyenesen elérhető Terminológiai Adatbázis-kezelő és a kereskedelmi forgalomban kapható *Kalcium quickTerm* előnyeit és hátrányait elemezzük és értékeljük.

Kulcsszavak: magyar terminológiastratégia, Magyar Nemzeti Terminológiai Adatbázis, szoftveres keretrendszer, szoftveres keretrendszer kiválasztásának szempontjai, adatbázis-kezelő szoftver

1 A kutatásokat a KRE BTK Terminológiai és Kommunikációs Kutatócsoport (TERMIK) keretében végeztük. A szerzők köszönik az MTA TMNP2023-1/2023. projektje támogatását, melyben az első szerző vezető tanácsadóként vesz részt.

FEASIBILITY OF THE HUNGARIAN TERMINOLOGY STRATEGY. SELECTION CRITERIA OF THE SOFTWARE FRAMEWORK FOR THE DESIGN OF THE HUNGARIAN NATIONAL TERMINOLOGY DATABASE

Abstract

Recently, the feasibility of the general Hungarian terminology strategy and at the same time the establishment of a legal and administrative terminology strategy have become important issues. One of the key tasks in the implementation of the terminology strategy is to collect terminological data, to process them from a terminological (linguistic and professional) point of view, and to publish them in terminological databases.

One of the tasks of the Hungarian terminology strategy is to design and build the Hungarian National Terminology Database. The paper aims to assess the publicly available frameworks that could be suitable for the management of the term base, their functionality, and their potential. In this study, we first consider the criteria for the selection of a software framework, and then analyze and evaluate the advantages and disadvantages of two database management software, *Terminologue*, a tool available free of charge, and *Kalcium quickTerm*, a commercially available system.

Keywords: Hungarian terminology strategy, Hungarian National Terminology Database, software framework, criteria for the selection of a software framework, database management software

1. Bevezetés

A magyar terminológiastratégia feladatai között szerepel a *Magyar Nemzeti Terminológiai Adatbázis* (MaNTA) megtervezése és felépítése.² Az adatbázis tervezése és előkészítése sok lépésből álló, nyelvészek és szakemberek együttműködését megkövetelő feladat. Várható haszna, hogy az adatbázis

2 Lásd FÓRIS Ágota – BÖLCSKEI Andrea: *Ajánlások a magyar terminológiastratégiához*. In: FÓRIS Ágota – BÖLCSKEI Andrea (szerk.): *Terminológiastratégiai kihívások a magyar nyelvterületen*. Budapest, L'Harmattan–OFFI Zrt., 140–164., <https://www.offi.hu/offi-akademia/kiadvanyok/terminologiastrategiai-kihivasok-a-magyar-nyelvterületen>, valamint PRÓSZÉKY Gábor (főszerk.) – FÓRIS Ágota – B. PAPP Eszter – BÖLCSKEI Andrea – LIPP Veronika (szerk.): *A magyar terminológiastratégia kialakítása. Zöld könyv*. Budapest, Nyelvtudományi Kutatóközpont, 2023, <https://doi.org/10.18135/term.2023> (2024. 08. 30.)

hozzájárul a magyar szakszókincs egyesítéséhez (harmonizációjához), ezzel megalapozza a későbbi egységesítési lehetőségeket, egyúttal lehetőséget nyújt a szaktudományos terminológiai kutatások kiszélesítéséhez.

Ahogy azt a jogi-közigazgatási terület terminológiastratégiájának kialakításával kapcsolatosan is írtuk,³ fontos volna a kormányzás területeit érintő jogi-közigazgatási terminológiastratégia kialakítása és az általános terminológiastratégiához történő kapcsolása. A megfelelő terminológiakezelés és dokumentumkezelés fontos elemei például az e-kormányzásnak, egyben a más nyelvek terminológiájával történő harmonizáció előfeltételei is. A jogi nyelvújítás, a jogi terminológia, a jogi szaknyelv, a közérthetőség kérdésköre szerves része a jogi-közigazgatási terminológiai munkáknak, ezen belül is a terminológiakezeléshez,⁴ konkrétan a Magyar Nemzeti Terminológiai Adatbázishoz a jövőben való lehetséges kapcsolódásához.

Egy másik tanulmányunkban⁵ az adatbázis tervezéséhez szükséges olyan általános kérdéseket mérlegeltük, mint hogy a megtervezendő és felépítendő adatbázisnak milyen technológiai követelményeknek kell megfelelnie és a működtetéséhez milyen informatikai háttérstruktúrára van szükség. Az általános megfontolásokon túl e kérdések részleteinek megvitatásához, a megfelelő szoftveres keretrendszer kiválasztásához kétféle alapinformációt mindenképpen ismernünk kell:

1. A leendő üzemeltető milyen technológiai és személyi háttérrel rendelkezik, például a szervere alkalmas-e az üzemeltetésre, illetve vannak-e olyan informatikusai és terminológusai (vagy erre felkészíthető kollégák), akik fel tudják építeni, de legalább hosszú távon tudják kezelni és működtetni a magyar nemzeti terminológiai adatbázist.
2. Milyen, nyilvános (akár nyílt forráskódú, akár kereskedelmi forgalomban kapható) keretrendszerek volnának alkalmasak a magyar nemzeti ter-

3 FÓRIS Ágota: A terminológiastratégia szempontjai a jogi-közigazgatási területen. *Glossa Iuridica*, 2023, 10 (3), 35–64.

4 RIXER Ádám: A magyar nyelv és magyar jogi műnyelv megújulása. *Glossa Iuridica*, 2014, 1 (1), 11–17.; RIXER Ádám: Szükség van-e Magyarországon jogi-közigazgatási szaknyelvújításra? Gondolatok egy kutatás margójára. *Glossa Iuridica*, 2023, 10 (3), 9–31. A jogi-közigazgatási terminológiastratégia kérdéseivel a KRE ÁJK Lőrincz Lajos Közjogi Kutatóműhelye 'Szükség van-e Magyarországon jogi-igazgatási szaknyelvújításra? Jog-igazgatási terminológiastratégia kialakítása' (2022–2024) című pályázata és az ennek keretében megjelent publikációk foglalkoztak, lásd a *Glossa Iuridica* folyóirat 2023, 10 (3) számában megjelent tanulmányokat.

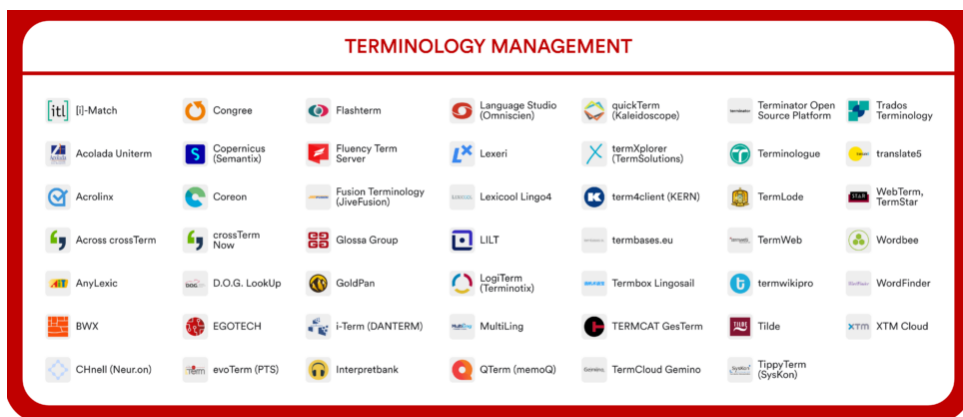
5 FÓRIS Ágota – SOMOGYI Zoltán – B. PAPP Eszter: Magyar nemzeti terminológiai adatbázis tervezése. Általános kérdések. *Alkalmazott Nyelvtudomány*, megjelenés alatt.

minológiai adatbázis létrehozására, kezelésére és a terminológiai adatok megjelenítésére a szerkesztők és a végfelhasználók számára, ezeknek milyen jellemzőik, tulajdonságaik és funkcióik vannak, és ezek milyen feltételekkel használhatók.

Jelen tanulmány célja a második kérdés vizsgálata, vagyis annak felmérése, hogy milyen olyan, nyilvánosan elérhető keretrendszerek léteznek, amelyek alkalmasak lehetnek a magyar nemzeti terminológiai adatbázis kezelésére, és ezeknek milyen funkcióik és milyen lehetőségeik érhetőek el.

2. A keretrendszer kiválasztásának szempontjai

A feladat tehát a terminológiai adatbázisok építéséhez és kezeléséhez felhasználható és elérhető keretrendszerek megvizsgálása és funkcióik, valamint lehetőségeik felmérése. Más nyelvtechnológiai alkalmazásokhoz hasonlóan, ma már a terminológiakezelő rendszerek széles palettája érhető el. A Nimdzi lokalizációs piackutatással foglalkozó szervezet évről évre megjelentet egy listát, amelyen az adott év legjelentősebb szoftvereit sorolja fel. Maga a lista több kategóriába sorolja a rendszereket, amelyek között a terminológiakezelő rendszerek is önállóan szerepelnek (lásd 1. ábra).



1. ábra. Terminológiakezelő rendszerek, 2023⁶

A 2023-as év listájában összesen 41 szoftver került a terminológiakezelő rendszerek kategóriájába. Jellegetek tekintve a keretrendszerek széles kör-

6 Forrás: NIMDZI Language Technology Atlas 2023, <https://www.nimdzi.com/language-technology-atlas/> (2024. 04. 03.)

ben mozognak, megtalálhatóak közöttük ingyenes, korlátozottan ingyenes, illetve fizetős modell szerint működők is. Ezek kapcsán fontos megjegyezni, hogy az 1. ábrán szereplő rendszerek többsége inkább a fordításorientált terminológiai munkát hivatott támogatni, funkcionális jellemzőik is inkább a fordítási munkafolyamat terminológiai részét helyezik előtérbe (adott esetben szorosan kapcsolódnak ugyanazon szoftvergyártó által kínált fordítástámogató eszközhöz is), ami miatt nem lehet automatikusan feltételezni, hogy megfelelő támogatást nyújthatnának olyan komplex terminológiai munka elvégzésére, aminek az eredménye egy nemzeti terminológiai adatbázis lehet.

A megfelelő szoftveres keretrendszer kiválasztásának első lépése *a nyilvánosan elérhető adatok felkutatása*. Ez elsősorban a szoftvergyártó cégek honlapját jelenti, ahol jó esetben a marketinganyagokon túl a videós vagy szöveges dokumentáció megismerésére is lehetőség nyílik. Ezek alapján már előzetes képet kaphatunk arról, hogy az általunk meghatározott igényekből mennyit tud kielégíteni a rendszer. A részletkérdések tisztázásához érdemes élő bemutatót is szervezni a technológiai szolgáltatóval, amely megbeszélésre célszerű a már összegyűjtött ismeretek alapján konkrét kérdésekkel érkezni, így segítve a végső döntést.

Természetesen evidens lehet olyan keretrendszert választani, amit – szem előtt tartva a nemzetközi szabványok előírásait és a jógyakorlatokat – kifejezetten terminológiai adatok és munkafolyamatok kezelésre hoztak létre,⁷ ugyanakkor a terminológiai projekt keretei számos korlátot és ezzel együtt lehetőségeket is magukban rejthetnek. A teljes projekt költségvetésében kulcsszerepet játszik a szoftveres háttér kialakításának és fenntartásának pénzügyi vonzata, ezért meg kell fontolni, hogy a) ingyenes, nyílt forráskódú vagy b) a piacról megvásárolható (kereskedelmi) rendszerre van-e lehetőség.

Ugyanakkor elképzelhető, hogy szóba kerülhetnek olyan keretrendszerek, amelyek alapvetően nem terminológiai adatok tárolására szolgálnak, de kellően rugalmasak lehetnek ahhoz, hogy megfelelő testreszabás mellett alkalmasak legyenek erre a célra. Itt olyan általános adatbázis rendszerekre gondolunk, mint például az Oracle által kínált APEX keretrendszer,⁸ amely low code, azaz túlságosan komoly programozási tudást nem igénylő rendszerként lehetőséget kínál bármilyen adatbázis és a hozzátartozó felhasználói felület

7 A terminológiai munkára vonatkozó szabványok áttekintését lásd SERMANN Eszter: *A terminológiai munkafolyamatok szempontjából releváns ISO-szabványok*. In: PRÓSZÉKY Gábor (főszerk.) et al. (2023) i. m. 279–300. <https://doi.org/10.18135/term.2023.14> (2024. 08. 30.)

8 Oracle APEX, <https://apex.oracle.com/en/> (2024. 04. 03.)

kialakítására. Ezeknek a platformoknak előnye, hogy már előre leprogramozott elemekből a felhasználók maguk is össze tudnak állítani jól működő rendszert. Érdemes kiemelni, hogy ezekben az esetekben elválhat a szerkesztői nézet, vagyis, hogy a) milyen módon tudnak a terminológusok adatokat bevinni az adatbázisba, illetve ezekben kereséseket/szűréseket végezni, tömegesen javítani, illetve b) az olvasói (végfelhasználói) nézet, mivel sokszor a meglévő háttér-adatbázisra valamilyen külső felhasználó számára is könnyen kezelhető felhasználói felületet kell külön ráépíteni. Ez a komplexitástól függően további programozási/webtervezési szaktudást igényelhet. A saját rendszerek megtervezése és kiépítése emellett mindig magában hordozza azt is, hogy a kialakított rendszert alapos és lehetőleg minden felhasználói szituációt lefedő tesztelésnek kell alávetni, illetve ilyen esetekben a karbantartási, valamint a funkcióbővítési feladatok is a projektben részt vevő technikai szakembereket terhelik. Összességében előfordulhat, hogy a saját rendszer fejlesztésével a piacon már elérhető, kiforrott rendszerek funkciókínálatának csak töredékét lehet megvalósítani, mindezt pedig ugyanolyan mértékű vagy akár jelentősebb anyagi és más erőforrás-ráfordítás mellett. Nilsson⁹ például azt írja, hogy azért fejlesztettek saját szoftvert a svéd *Rikstermbanken* terminológiai adatbázis kezeléséhez, mert 2006-ban nem állt rendelkezésre olyan teljesítményű szoftver, amely megfelelt volna számukra.

A nulláról való felépítés mellett felmerülhet még egy *meglévő technológia adaptálása is terminológiai célokra*. Ha a magyar példát nézzük, akkor szóba jöhet akár a nagyszótári projektben is használt keretrendszerek felhasználása. A magyar nyelv nagyszótára¹⁰ szerkesztési munkálatai egy XML-szerkesztőben (XMetal)¹¹ történnek, illetve egy nyílt forráskódú XML-alapú adatbázisban (eXist-db)¹² tárolódnak. Annak ellenére, hogy a *Nagyszótár* szemlélete más, mint egy terminológiai adatbázisé (szemasziológiai vs. onomasziológiai szemlélet), a szerkesztési technológia adaptálható lenne a terminológiai munkákra is, hiszen számos terminológiai fájl típus XML-alapú, köztük a legismertebb, a TBX is, amelyről önálló szabvány is született (ISO 30042:2019).¹³ Az adaptál-

9 NILSSON, Henrik: The realisation of a national term bank – how and why? In: Proceedings of the ELETO – 7th Conference 'Hellenic Language and Terminology' Athens, Greece, 22-24 October 2009. https://www.eleto.gr/download/Conferences/7th%20Conference/7th_26-26-NilssonHenrik_Paper_V03.pdf (2024. 04. 03.), 6.

10 A magyar nyelv nagyszótára [Nagyszótár], <https://nagyszotar.nytud.hu/index.html> (2024. 04. 03.)

11 XMetal, <https://xmetal.com/> (2024. 04. 03.)

12 eXist-db. <http://www.exist-db.org/exist/apps/homepage/index.html> (2024. 04. 03.)

13 ISO 30042:2019 Management of terminology resources. TermBase eXchange (TBX).

hatóság mellett a szerkesztés technikai gyakorlatának tapasztalatait is meg lehetne osztani a szótáron már régóta dolgozó kollégák és a terminológusok között. Természetesen ezt az utat választva sem lehet elkerülni a további fejlesztések szükségességét: az adaptálás után még mindig ki kell alakítani egy olyan keresőfelületet, amelyet a végfelhasználók használhatnak, illetve valamilyen megoldást kell találni a munkafolyamatok szoftveres központosítására, valamint ezek kisebb-nagyobb mértékű automatizálására is.

Akármilyen keretrendszert is veszünk fontolóra, mielőtt végleg elköteleződnénk egyik vagy másik mellett, *alapos tesztelésnek kell alávetnünk a rendszereket*. Ennek alapját mindazok az igények és megfontolások képezik, amelyeket az előzetes tervezés és az adott rendszer által nyújtott lehetőségek figyelembevételével határozzunk meg. A tesztelés során érdemes nagyobb mennyiségű adattal feltölteni a rendszert. Ennek során egyrészt kitapasztalhatjuk szerkesztői szempontból, hogy mennyire intuitív a felület, milyen mélységig lehet finomhangolni a rendszert, illetve mennyire nehézkes az adatbevitel. Mindenképpen érdemes a másik oldalról, vagyis a végfelhasználó oldaláról is tesztelni a rendszereket, itt is elsősorban az intuitivitás kérdése kap hangsúlyos szerepet. Hiszen készíthetünk ugyan nagyon részletes felhasználói dokumentációt, a legtöbb ember mégis inkább saját maga szereti felfedezni a különböző rendszereket, és a felhasználói élményben alapvető nyomot hagy az, hogyha a felhasználói felület elsőre nem könnyen és egyértelműen használható. Fontos megnézni azt is, hogy milyen módon jelennek meg az egyes adatkategóriák; itt érdemes kitérni nem csak a szöveg elrendezésére, hanem hogy a színek, formák és egyéb kiemelési módok milyen módon segítik a tájékozódást, milyen lehetőségeket kínál a rendszer a végfelhasználók figyelmének irányítására. Összességében el lehet mondani, hogy minél hosszabb a tesztelési szakasz, annál inkább megismerjük a szoftver működését és annak korlátait, amely mind segítségünkre lesz a végső rendszer kiválasztásában. Mindennek alapvető szerepe van az egész projekt sikerességében, hiszen mind anyagi mind más erőforrás-ráfordítás tekintetében nagyon nehéz projekt közben keretrendszert váltani, illetve meghatározza a terminológiai munka hatékonyságát is. A cél az adott körülmények között a leggyorsabb és leghatékonyabb módon a lehető legpontosabb adatok bevitele és mindezek megfelelő formában történő tárolása a végfelhasználó számára.

3. Konkrét adatbázis-kezelő rendszerek vizsgálata

Ebben a részben két rendszert vizsgálunk. A cél annak meghatározása, hogy a 2. pontban ismertetett igényeknek mennyire képesek megfelelni. A termékínalat két széléről választottunk: az egyik rendszer a *Terminologue*, a másik pedig a *Kalcium quickTerm*. Az elemzés nem terjed ki a két rendszer szisztematikus összehasonlítására, sokkal inkább előnyeik és hátrányaik bemutatására. A rendelkezésre álló források és lehetőségek is eltérőek: a) a *Terminologue* esetén csak egy rövid dokumentáció áll rendelkezésre, de mivel ingyenes, könnyen lehet tesztelni, addig b) a *quickTerm* esetén bővebb felhasználói dokumentáció és videós anyagok érhetők el, viszont nem volt alkalmunk élesben tesztelni a rendszert.

3.1. *Terminologue*

A *Terminologue*¹⁴ terminológiamenedzsment rendszer egy felhőalapú, ingyenes és nyílt forráskódú keretrendszer, amelyet a Dublini Egyetem (Dublin City University) Gaois Research Group kutatócsoportja fejlesztett elsősorban az Ír Nemzeti Terminológiai Adatbázishoz (*National Terminology Database for Irish*, NTD), ám ezt később mindenki számára hozzáférhetővé tették.¹⁵ Maga a keretrendszer futtatható a Dublini Egyetem szerverein is, ebben az esetben onnan is lesz elérhető az adatbázis, ugyanakkor a fejlesztők a teljes programot megosztották a GitHubon,¹⁶ így bárki számára biztosították, hogy a keretrendszert a saját szerverére telepíthesse és onnan tehesse elérhetővé.

A felhasználói felület nyelve ugyan magyarul nem érhető el, de a fejlesztők megadják a lehetőséget arra, hogy bárki lokalizálhassa a szoftverelemeket a saját nyelvére. Ez meg is történt azon 14 nyelv esetében, amelyekben a rendszer jelenleg elérhető, mindemellett pedig a honlap feltünteteti a fordítói

14 Terminologue, <https://www.terminologue.org/> (2024. 04. 03.)

15 MECHURA, M. – Ó RAGHALLAIGH, B.: *Introducing Terminologue: a cloud-based, open-source terminology management tool*. In: *Lexicography for Inclusion: Proceedings of the 19th EURALEX International Congress, 7-9 September 2021, Alexandroupolis, Democritus University of Thrace*. 2021, Vol. 2, 797–800. https://euralex.org/wp-content/themes/euralex/proceedings/Euralex%202020-2021/EURALEX2020-2021_Vol2-p797-800.pdf (2024. 04. 03.), valamint BHREATHNACH, Ú., Ó CLEIRCÍN, G.: *Terminologue as a teaching resource for terminology - assessing the user experience*. Paper presented at the 11th EAFT Summit, 2023. Terminology – Innovation and Sustainability. Barcelona (November 16-17, 2023.)

16 GitHub, <https://github.com/gaois/terminologue> (2024. 04. 03.)

csoportok vagy intézmények nevét is. A szerkesztési felületre vonatkozóan ez az adatbázis lehetővé teszi azt is, hogy a különböző adatkategóriák nevét is az adott nyelven lehessen elérni, így nyelvi felhasználási szempontból minden igényt ki tud elégíteni.

A tesztelés során létrehoztunk egy teszt-adatbázist a nyilvánosan elérhető szerveren. Az adatbázis bárki számára elérhető a <https://www.terminologie.org/MaNTA/> címen.¹⁷ Az URL-címből jól látszik, hogy az adatbázis nem saját szerveren működik; ennek kapcsán érdemes megjegyezni, hogy a keretrendszer lehetőséget biztosít az URL-cím testreszabására, így segítve a felhasználót a könnyebb tájékozódásban. A tesztelés során 24 bejegyzést vittünk be, ebből 23 'A magyar nyelv és a szakmai kultúra. Magyar nyelvészeti kutatások a tartalomfejlesztés és a dokumentáció területén'¹⁸ című kutatási projekt eddig még nem publikált terminusgyűjteményéből származik, a maradék 1 bejegyzés pedig a *DictionELI*¹⁹ lézerfizikai terminológiai szótárból.

Végfelhasználói szempontból könnyű az adatbázisban keresést végezni. A nyitólapon található szöveges beviteli mezővel lehet szóalapú kereséseket futtatni. A fejlesztők számításba vették, hogy a felhasználók nem mindig pontosan adják meg a keresendő terminusokat, ezért olyan mechanizmusokat építettek a keresési algoritmusba, amelyek ezt a jelenséget is kompenzálják.²⁰ Alapvetően elmondható, hogy csak a terminusokra lehet keresni, tehát ha valaki olyan keresőszót ír be, ami pl. a definícióban fordul elő, akkor a rendszer nem hoz találatot. Ennek az is a következménye, hogy a rövidített szóalakokat külön terminusként kell hozzáadni, nincs lehetőség más adatkategóriát használni, ha meg szeretnénk jeleníteni a keresési találatok között. A találatok között szűrni ugyan nem lehet, viszont lehetőség van a megjelenített bejegyzéseken a domén, illetve aldomén kategóriára kattintva kilistázni, hogy mi tartozik az adott kategóriához. Ezek alapján megállapítható, hogy a végfelhasználó nem tud összetett kereséseket futtatni, vagyis nem tudja a szöveges és egyéb keresési módok kombinációjával finomhangolni a keresés eredményét. A szer-

17 Teszt-adatbázis, <https://www.terminologie.org/MaNTA/> (2024. 04. 03.); készítette: SOMOGYI Zoltán.

18 A KRE BTK Terminológiai és Kommunikációs Kutatócsoport (TERMIK) 2018–2021 között folyó, a KRE BTK által támogatott projektje. Összefoglaló kötet: FÓRIS Ágota – BÖLCSKEI Andrea (eds.): *Linguistic Research in the Fields of Content Development and Documentation*. Budapest–Paris, KRE – L'Harmattan, 2021.

19 FÓRIS Ágota – B. PAPP Eszter – SERMANN Eszter (szerk.): *DictionELI. Lézerterminológiai szótár*. Szeged, Szegedi Tudományegyetem, 2015., <http://dictioneli.stepp.hu/> (2024. 04. 03.)

20 MĚCHURA – Ó RAGHALLAIGH (2021) i. m.

kesztői felület azonban lehetőséget ad ezekre, így áthidaló megoldás lehet a végfelhasználónak egy olyan felhasználói fiókot létrehozni, amivel csak olvasni lehet (létezik ilyen jogkör). Ugyanakkor minden bizonnyal előfordulhat, hogy egyszerre csak egy felhasználó használhatja ezzel a fiókkal az adatbázist, ami egy nemzeti terminológiai adatbázis esetében elképzelhetetlen.

A bejegyzések megjelenítésével kapcsolatban elmondható, hogy a végfelhasználói nézet és a szerkesztői nézet szinte teljesen ugyanúgy néz ki. Ez azt is jelenti, hogy a keretrendszerben nem lehet szabályozni, hogy mely adatkategóriákat lássák a végfelhasználók, és melyeket ne. Emellett a végfelhasználók számára viszont nem láthatók az adminisztratív adatkategóriák, vagyis, hogy az adott bejegyzést ki és mikor módosította utoljára. A szerkesztői felületen, a bejegyzéseknél ugyan meg lehet nézni a bejegyzés előzményeit (ki, mikor és mit változtatott), sőt még vissza is lehet állítani a korábbi állapotokat, sajnos azonban ez a végfelhasználó számára nem elérhető opció. Ez a csoport csak akkor láthatná ezeket az adatokat, ha a fentiek szerint csak olvasási joggal rendelkező fiókkal beléphetne a szerkesztői oldalra, viszont ez esetben nem lehet azt korlátozni, hogy csak a legutolsó módosítási dátumot lássa, hanem láthatná az adott bejegyzés teljes előzményét, ami viszont nem lenne releváns számára. Ugyanakkor a szerkesztői felületen is csak a változtatást elvégző felhasználó e-mail-címe látható, a neve nem, tehát ha nem elég beszédes e-mail-címe van valakinek, nem tudható, hogy ki végezte a változtatást. Ennek a végfelhasználó szempontjából nincs jelentősége, a szerkesztői felhasználók és a terminológiai munkacsoport számára viszont kulcsfontosságú volna az információ. Erre a problémára áthidaló megoldás lehet egy külön megjegyzés típusú adatkategóriát létrehozni, és abban kézzel vezetni a módosítás időpontját, illetve a módosító felhasználót, de ez nagy terhet róna a szerkesztőkre.

A megjelenítési szempontok másik fő kérdése az is, hogy az egyes nyelveket miképpen lehet megjeleníteni. A *Terminologue* rendszeréről elmondható, hogy egy-, illetve kétnyelvű adatokat áttekinthetően meg tud jeleníteni, ugyanakkor, ha három vagy akár több nyelv (vagy nyelvváltozat) kerül az adatbázisba, akkor az egyes bejegyzések egymás mellé, oszlopszerűen kerülnek, a tördelés nagyon szűkre szabott lesz, ez pedig nehezíti a végfelhasználó számára az információ befogadását.

Az elrendezésen túl, ha lemegyünk az egyes adatkategóriák szintjére, akkor megfigyelhető, hogy a hipervivatkozásokot tartalmazó adatmezőkben maguk a hivatkozások sima szöveggént szerepelnek, vagyis nem kattinthatók. Ugyanakkor a rendszer lehetővé teszi a bejegyzések összekapcsolását,

amelyet a végfelhasználók a kapcsolódó terminusok kategória alatt láthatnak, és ezeknél működik a kattinthatóság, ami nagyban hozzájárul ahhoz, hogy a végfelhasználók összefüggő képet alkothassanak az adott domén vagy aldómén fogalmi hálózatáról. A kapcsolódó terminusoknál azonban a kapcsolat típusát nem lehet meghatározni, illetve gráf megjelenítésére sincs lehetőség.

A szerkesztési felületet vizsgálva az látszik, hogy az adatbázis valamelyest eltér a nemzetközi szabványoktól és ajánlásoktól, hiszen nem különíti el a szerkezetben a bejegyzés, a nyelvi indexálás és a terminus szintjét.²¹ Ennek igazán zavaró gyakorlati jelentősége nincs, hiszen bizonyos alap-adatkategóriák, úgymint a terminus, a definíció, a kontextus, a doménbarendeozhetőség, és egyéb, egyénileg meghatározható adatkategóriák elérhetők, illetve létrehozhatók a felületen. Többségük már eleve beépített a rendszerben, tehát a *Terminologue*-ban nem kell és nem is lehet teljesen egyéni adatkategóriákat és adatstruktúrát létrehozni, hanem ezek mind adottak a rendszeren belül. Ugyanakkor elmondható, hogy az egyes adatkategóriák finomhangolására számos lehetőség van. A terminusok vonatkozásában a szokásos grammatikai információkon túl (pl. szófaj megadása) további morfológiai jellemzők is rögzíthetők. Ezen adatkategóriák elemei teljes mértékben személyre szabhatók a rendszer adminisztrációs felületén, sőt, az egyes elnevezések bármikor módosíthatók anélkül, hogy bármilyen exportálási-importálási vagy manuális szerkesztési műveletet el kellene végezni. Ugyanakkor nincs lehetőség semmilyen tömeges szerkesztési művelet elvégzésére, vagyis a bejegyzéseket nem lehet egy lépésben átkategorizálni, csupán a kategóriák elnevezését lehet megváltoztatni, ám ez kihat az adott kategória alá tartozó összes bejegyzésre.

Rátérve az adatkategóriák megjelenítésére, kissé furcsa, hogy van ugyan lehetőség a definíció és a kontextus megadására, de ezek vizuálisan mégis szokatlan módon válnak el: a kontextus félkövér, míg a definíció nem. A megjelenítés e tekintetben sem módosítható, ez a beállítás a rendszeren belül van rögzítve. A definícióhoz és a kontextushoz egyaránt lehet forrást megadni, ezek szürkével jelennek meg az adott adatkategória alatt, de nem lehet megjeleníteni, hogy mit lát a felhasználó, azaz nem lehet odaírni, hogy ez a definíció forrása vagy a kontextus forrása. Létezik egy megjegyzés típusú adatkategória, amit bárhogy el lehet nevezni és amit lehet(ne) használni erre a célra is, de akkor a háttérben, adatbázisszinten lesz nem megfelelő a kategória (vagyis programszinten az adott adatkategória megjegyzésként és nem a definíció vagy a kontextus forrásaként tárolódik, ami hatással van a rendszerből kinyerhető fájlformátumokra is).

21 Uo.

Mind a definíció, mind a kontextus forrása szabad szöveges beviteli mező, ugyanakkor az adatbázis ezeket külön adattáblában kezeli, tehát van olyan felület, ahol az adott terminológiai adatbázishoz tartozó összes forrás áttekinthető, az adatokban módosítás végezhető, ami átöröklődik az összes olyan bejegyzésre, ahol az adott forrást felhasználták. Így a meglévő adatokat újra fel lehet használni, ha egy adott adat ugyanabból a forrásból származik, a szoftver lehetőséget ad kiválasztani a meglévő bibliográfiai adatokat egy legördülő listából. Annak ellenére, hogy ez előnyösnek tűnhet, számos problémát vet fel. Bizonyos adatszám felett nehezen átláthatóak az ilyen listák, még akkor is, ha a legördülő menüben ábécé sorrendben vannak feltüntetve. Különösen igaz ez akkor, ha a terminológiai adatbázis több domént és aldomént foglal magába – mint a tervek szerint a magyar nemzeti terminológiai adatbázis.

A szerkesztői felület már több lehetőséget biztosít a keresésre, illetve a szűrésre, itt szöveges keresés is futtatható a terminusok, definíciók, kontextusok, megjegyzések, gyűjtemények, valamint a szerkesztői felhasználói megjegyzések között is. A keresés során lehet 'tartalmazza', illetve 'nem tartalmazza' típusú szűrést is végezni. Mind a szűrést, mind pedig az adott bejegyzés életciklusának adott pillanatban meglévő állapotát segítik az adminisztratív kategóriák, amelyek szintén az adatbázis alapját képezik. Ennél a kategóriánál három érték lehetséges: *ellenőrzöttségi státusz (checking status)*, *közzétételi státusz (publishing status)*, valamint *piszkozat státusz (drafting status)*. Ez egyrészt segítheti a terminológiai munkafolyamatot is (hiszen maga a rendszer nem rendelkezik olyan összetett projekt- vagy feladatkezelési modullal, ami lehetővé tenné a feladatok kiosztását és nyomon követését), illetve szabályozni tudja, hogy mely bejegyzések vannak már olyan állapotban, hogy azokat a végfelhasználó is láthassa. Ugyanakkor hozzá kell tenni, hogy ezen státuszokat a szerkesztő mindig manuálisan állítja be minden egyes bejegyzésnél, erre nincs semmilyen automatizálási mód.

A fenti hiányosságok, főleg az, hogy nem lehet felhasználók szerint szűrni az adatot, egy többtagú terminológusi csoportnál megnehezítik a projekt haladásának nyomon követését, illetve az adatok ellenőrzését. Erre rész megoldást kínál a bejegyzés különböző státuszainak változtathatósága, és az arra való szűrés lehetősége, de ez még mindig messze van a munkafolyamatok technológiai támogatásától. A korábbiakban már szó volt róla, hogy a rendszer nem ad lehetőséget semmilyen tömeges szerkesztésre, ez pedig a keresési/szűrés lehetőségek korlátai mellett, egy bizonyos bejegyzésszám felett nagyon megnehezíti az új adatok bevitelét, illetve az adatbázis szerkesztési szintű karbantartását. Nincs lehetőség keresés/csere-típusú műveletet elvégezni.

Jóllehet bizonyos adatkategóriákat lehet módosítani globálisan (leginkább az adatkategóriák címkéjét, de akár konkrét adatmezőket is, pl. bibliográfiai adatok), ám ez nem mindenhol valósítható meg. Ha tömegesen szeretnénk megváltoztatni mondjuk a bejegyzések státuszát vagy a bejegyzések domén/aldomén-besorolását, akkor azt csak manuálisan, egyesével lehet megtenni. Nagyon korlátozott a lehetőség arra, hogy összekapcsoljunk bejegyzéseket, csupán egy sima 'related term' nevű, előre beépített mező van. A rendszer nem kínál még előre definiált kapcsolattípusokat sem, illetve nem támogatja az egyéni kapcsolatok létrehozását sem. Ezáltal a bejegyzések egymáshoz való viszonyát csak szövegszinten lehet megvalósítani, nincs mód a vizualizációra, emellett nem lehet a rendszerből olyan fájlformátumot kinyerni, amelyet más, akár ontológiakezelő rendszerekben tovább fel lehetne dolgozni, így az adatbázis nem alkalmas a gépi modulok betanítására vagy az azokhoz való felhasználásra.

Még tágabb technológiai szempontból, a keretrendszer pozitívumaként emelendők ki a felhasználókezelési opciók. Egy ötfokozatú skálán lehet megadni, hogy egy különálló fiókkal rendelkező felhasználó milyen szintű hozzáféréssel rendelkezzen a sima olvasási hozzáféréstől a minden szerkesztési és az adatbázis struktúráját is módosítani képes jogkörig.

További fontos szempont még az exportálhatóság és az adatok más rendszerekben való feldolgozhatóságának vagy más rendszerekbe való átültetésének kérdése is. A *Terminologue* keretrendszeréből háromféle formátumban nyerhetők ki az adatok. Az első a TBX, amelyről elmondható, hogy a rendszer által generált változat megfelel a fájlformátumra vonatkozó szabványnak.²² A megfelelőségről a TBX Spyglass²³ nevű webes programmal lehet meggyőződni, amelynek segítségével meg lehet állapítani, hogy a feltöltött fájl megfelel-e a szabvány által meghatározott specifikációknak, illetve pontosan melyik változatának, a 2-es vagy a 3-as verzióknak (a *Terminologue* által generált fájl az utóbbi verzióhoz tartozik). Emellett létezik egy másik webes alkalmazás is (TBX Viewer),²⁴ amelynek segítségével azt ellenőrizhetjük, hogy az exportált fájl pontosan milyen adatokat tartalmaz. Maga a TBX mint fájlformátum XML-alapú, tehát az emberek számára is olvasható eredeti formában, ugyanakkor a TBX Viewerhez hasonló alkalmazással még inkább áttekinthető lesz a fájl tényleges tartalma. Az exportált TBX vizsgálatok az állapítható meg, hogy az sajnos nem tartalmaz minden adatot vagy olyan módon tárolja

22 ISO 30042:2019 i. m.

23 TBX Spyglass, <https://www.tbxinfo.net/tbx-spyglass/> (2024. 04. 03.)

24 TBX Viewer, <https://viewer.tbxinfo.net/> (2024. 04. 03.)

ezeket, hogy később nehezebben lehet feldolgozni más rendszerben. Erre jó példa a doménok és aldoménok elválasztása, amely lúdláb karakterrel történik (Domén » Aldomén formában), és mindkettő a subjectField elem alá van besorolva a fájlban belül. Emellett az is látható, hogy amint az olvasói felületen, úgy a TBX exportban sem látszanak az adminisztratív metaadatok, vagyis, hogy az adott bejegyzést ki és mikor módosította utoljára. Szintén hiányoznak a kapcsolódó terminusok is, tehát az előző adatkategóriával együtt az adatbázis ezen része is elvész exportáláskor.

A második fájlformátum, amelyet a rendszer exportálni képes, az a sima, tabulátorral elválasztott szövegfájl. Ebben az esetben minden adatmező fel van sorolva az első sorban és amelyekben található adat az adott bejegyzésnél, azt a rendszer kiírja a fájlba. Olyan adatmezők esetén, ahol a mezőben található sortörés, ebben az exportálási formátumban '[LINEBREAK]' karaktersort ír a program, ami megkönnyítheti a további feldolgozást vagy egy másik keretrendszerbe való átültetést. Némi nehézséget okozhat, hogy a szófaji címkék a terminus mellett, szóközzel elválasztva jelennek meg, nem pedig önálló oszlopként. A további adatkategóriák vonatkozásában erre a formátumra is érvényesek a TBX formátumnál említett összevonások és hiányzó adatkategóriák, tehát összességében csak kis mértékben tér el ez a formátum az előzőtől.

A harmadik fájlformátum az SQLite fájl, ami gyakorlatilag a teljes relációs adatbázis egy fájlba való összesítését jelenti. Ennek a formátumnak az egyik előnye, hogy minden adatot tartalmaz, amelyet a szerkesztési felületen a felhasználók bevitték, a konkrét adatkategóriáktól egészen az olyan metaadatokig, mint a bejegyzés életciklusához tartozó adatok. A hátránya éppen az előnyéből következik: az adatok kinyeréséhez, illetve további feldolgozásához komolyabb adatbázis-kezelési tudás szükséges (mindenelőtt az SQL lekérdezési nyelv ismerete). Ennek kapcsán azonban azt is meg kell állapítani, hogy sajnos a dokumentáció nem tartalmazza, hogy az adatbázis milyen táblákból áll, melyik kulcsok kötik össze a táblákat, milyen módon lehet lekérdezéseket futtatni, éppen ezért ennek feltárása és a hipotézisek tesztelése további erőfeszítést és mélyebb szaktudást igényel. Az adatbázison belül maguk az adatok JSON fájlformátumban tárolódnak,²⁵ amelyben a nem ASCII-karaktereket a rendszer külön is kódolja (pl. ű helyett \u0171 kódot használ), ami újabb akadályt gördít az adatbázis további feldolgozhatósága elé.

Infrastrukturális szempontból nem ismert a rendszer terhelhetősége. Bhreathnach és Ó Cleircín arról számolnak be,²⁶ hogy jelenleg az ír nemzeti

25 MĚCHURA – Ó RAGHALLAIGH (2021) i. m.

26 BHREATHNACH – Ó CLEIRCÍN (2023) i. m.

terminológiai adatbázis több mint 186 000 bejegyzést tartalmaz, valamint a *Terminologue* rendszert a terminológia oktatásában számos egyetemi projektben használják, így több mint 2500 regisztrált felhasználó, több mint 3000 adatbázisban dolgozik vagy dolgozott korábban. Az azonban nem tudható, hogy igazán nagy bejegyzésszámnál, valamint több felhasználó együttes munkájánál mennyire válik nehézkesé a keresés és a szerkesztés mind a végfelhasználói, mind pedig a szerkesztői oldalon. Mindez vajon csak a szoftvert futtató szerver teljesítményén múlik vagy vannak kapacitásbeli korlátai a mögöttes adatbázis-technológiának is? Sajnos az előzetes vizsgálati fázisban ezeket nem lehet szimulálni.

Összességében elmondható, hogy a *Terminologue* által nyújtott funkciók és lehetőségek meglehetősen jók, ha számításba vesszük azt a tényt, hogy a rendszer ingyenesen hozzáférhető és nyílt forráskódú. Ugyanakkor a fentebb említett korlátok miatt nem biztos, hogy ez a legalkalmasabb keretrendszer egy komplex nemzeti terminológiai projekt technológiai megvalósítására.

3.2. *Kalcium quickTerm*

A *Kalcium quickTerm*²⁷ termék mögött az ausztriai Kaleidoscope nevű cég áll. Ez a termék jellegét tekintve két szempontból is a *Terminologue* ellenpárjaként fogható fel: egyrészt ez egy kereskedelmi, tehát nem szabadon hozzáférhető, nem nyílt forráskódú rendszer, másfelől pedig kimondottan a vállalati terminológiamenedzsment igényeire szabták, igyekezve leképezni technológiailag annak minden aspektusát és komplexitását. A terméknev első fele, a Kalcium, egy különálló platformra utal, amelyben a *quickTerm* mint termék csupán egy a sok közül: a cég többek között kínál még lokalizációs minőségellenőrzési és kérdéskezelési megoldásokat is. Maga a rendszer kezdetben a *MultiTerm* programra építve, ahhoz kapcsolódó munkafolyamatkezelést kínált elsősorban, azonban mára elmondható, hogy a kínálat egy *teljesen platformfüggetlen terminológiamenedzsment rendszerré* nőtte ki magát, amely számos szoftverhez képes kapcsolódni a fordítástámogató és szakszövegíró alkalmazásoktól kezdve az egyszerű szövegszerkesztő programokig, így segítve a felhasználókat az egy- és többnyelvű terminológiához való igazodáshoz minden felhasználási fázisban.

A *Terminologue*-gal szemben a *quickTerm* elemzése során nem tudtunk konkrét tesztelési fázisra támaszkodni, csupán a nyilvánosan elérhető ada-

27 Kalcium quickTerm, <https://kaleidoscope.at/en/products/quickterm/> (2024. 04. 03.)

tokból, elsősorban az egyes programverziókat bemutató webináriumokból²⁸ tudtunk tájékozódni. Emiatt a relevánsabb funkciók bemutatása nem lehet mindenre kiterjedő.

A rendszer felhasználói felületének választható nyelvei között nem szerepel a magyar, de az elérhető források szerint a cég nyitott a különböző ügyféligenyek mentén bármilyen nyelvet hozzáadni a listához. A nyelv és fordítás kérdésében érdemes megjegyezni, hogy a rendszer lehetőséget biztosít az egyes adatkategóriák, valamint a szerkesztői és végfelhasználókat segítő súgószövegek lokalizálására, mindezt pedig nem muszáj a rendszeren belül végezni, hanem egy ennek a feladatnak az elvégzésére alkalmas fájlformátumba (RESX) való exportálással azon kívül is el lehet végezni, akár fordítástámogató programokkal is. Ez a formátum azért is előnyös, mert nagy mennyiségű adatkategória, valamint az azokon belüli értékek nagy száma esetén biztosítani lehet a következetes fordítást gépi eszközökkel. A szoftver képes a bejegyzések közötti kapcsolatok vizuális megjelenítésére kapcsolati gráfok formájában. Ennek lokalizációs szempontból az a jelentősége, hogy az egyes terminusok közötti kapcsolat, valamint az ezeket magyarázó kiegészítő szövegek is fordíthatók, így valósulhat meg, hogy ha a felhasználó nyelvet vált, akkor valóban minden egyes szoftverelem nyelve megváltozik.

A végfelhasználói élmény egyik fő alkotóeleme az adatbázis kezdőképernyője. A *quickTerm* lehetőséget ad arra, hogy különböző elemtípusokkal lehessen gazdagítani ezt az oldalt. Az általános szöveges részek mellett, amelyek lehetnek például üdvözlő szövegek vagy az adatbázisról szóló rövid információk, lehetőség van kiemelt bejegyzéseket is megjeleníteni az oldalon, akár időszakosan is (például a hét vagy a hónap terminusa formájában). Ezek a funkciók hozzájárulnak ahhoz, hogy az adatbázis közelebb kerüljön a felhasználóhoz. A felhasználói élmény további figyelemreméltó eleme az egyéni testreszabhatóság, amely megjelenik a sötét témára való váltás lehetőségében vagy éppenséggel a magas kontrasztú szín- és elemösszetételben is, amely az akadálymentesség egyik kulcsfontosságú eleme (vö. a webes tartalmak akadálymentességéről szóló *Web Content Accessibility Guidelines*²⁹ nemzetközi útmutatóval).

A keresés tekintetében nagyon pontosan lehet szabályozni, hogy a végfelhasználó milyen keresést tudjon elvégezni, illetve ennek a konkrét fel-

28 Kalcium quickTerm webináriumok, <https://www.youtube.com/watch?v=VqMUSOoHg3o-amp;list=PLw1ECjgQB6sXPKvqTZAnewc6zKMQ5qhF&pp=iAQB> (2024. 04. 03.)

29 Web Content Accessibility Guidelines (WCAG) 2.0. <https://www.w3.org/TR/2008/REC-WCAG20-20081211/> (2024. 04. 03.)

használói felületen megtalálható elemeit (gombok, legördülő listák és egyéb szabályozók) működését és megjelenítését a háttérben finomhangolni lehet. Az egyszerű szöveges bevétel során a rendszer prediktív módon igyekszik már egy legördülő lista formájában megjeleníteni a lehetséges keresőszavakat, ami segíti a felhasználót abban is, hogy egy gyors képet kapjon az adatbázisban megtalálható, a keresési kifejezés töredékével megegyező bejegyzésekről is. A szöveges keresést az egyéb kategóriákkal való kiegészítés teszi még pontosabbá. A kategóriák megjelenítési módja kapcsán kiemelendő a taxonimikus keresés lehetősége, amely a gyakorlatban azt jelenti, hogy az egyes fő- és alkategóriák egymásból lenyíló ablakokban jeleníthetők meg.

A bejegyzések megjelenítésének módja is teljes mértékben testreszabható. Stíluslapok segítségével meghatározható, hogy milyen felhasználói csoport milyen és hány darab adatkategóriát lásson a bejegyzésen belül egy adott keresés eredményeként. Azonban nem csak a kategóriák száma és elrendezése választható meg, hanem a tájékozódást segítve színekkel is ki lehet emelni a jóváhagyott, javasolt vagy éppen tiltott terminusokat (ez az adatbázis preskriptív vagy deskriptív jellegétől függ). Természetesen a színek használata is meghatározó tényező lehet az akadálymentesítés vonatkozásában, ezért a rendszer lehetőséget ad ikonok feltöltésére is, amelyekkel lehet jelezni a preferált vagy kerülendő terminusokat.

Felhasználói szempontból kiemelendő funkció a környezeti súgók alkalmazhatóságának lehetősége. Ez a gyakorlatban azt jelenti, hogy ha a felhasználó az egérmutatóval egy adatkategória vagy más gomb fölé megy, akkor megjelenik egy rövid súgószöveg az adott adatkategóriához vagy funkcióhoz kapcsolódó használati vagy egyéb információkról. Ehhez kapcsolódó kiegészítő funkció még az is, hogy ezt a szöveget HTML szerkesztőben lehet elkészíteni, vagyis lehet bele tenni kiemeléseket, strukturálni lehet a szöveget, illetve további kiegészítendő hiperhivatkozásokat lehet elhelyezni benne. Ezen mezők megéléte pozitívan tudja befolyásolni a felhasználói élményt, hiszen a felhasználó nincs arra kényszerítve, hogy a hosszú felhasználói dokumentációt kelljen böngésznie, ha szeretne több információt megtudni az adott interfészlelemről. Természetesen ezen a szövegek is mind lokalizálhatók.

A szerkesztő felhasználói szempontok közül az első és legfontosabb az, hogy a rendszer lehetőséget ad egyéni nyelvváltozatok felvételére is az adatbázisba. Ez a piacon lévő terminológiamenedzsment rendszerek többségében nem elérhető, csak az előre meghatározott nyelvlistánból lehet választani. Egy magyar nemzeti terminológiai adatbázisnál azonban ez a lehetőség elengedhetetlen, hiszen a terminológiai projekt egyik fő célkitűzése lehet az

országhatárokon túli vagy éppen az azokon belüli nyelvek és nyelvváltozatok rögzítése az adatbázisba. Vizuális kiegészítésként az egyénileg létrehozott nyelveket egyéni zászlóikonokkal is ki lehet egészíteni.

Az adatkategóriák és a terminológiai adatbázis módosításával kapcsolatban további tudnivaló, hogy a rendszer lehetőséget ad az adatmezők átnevezésére, valamint a korábban meghatározott listaelemek kiegészítésére, módosítására, törlésére, vagy inaktíválásra (ez utóbbi azt jelenti, hogy az adatbázis-szerkezet nézetében jelen van az érték, de a szerkesztő felhasználó számára nem jelenik meg választhatóként). Az adatkategóriák egymással való viszonyát is lehet szabályozni olyan módon, hogy fölé- és alárendelt értékpárokat lehet meghatározni, vagyis például adott doménhoz csak bizonyos számú aldomént lehet választani. A másik, szerkesztői felhasználókat segítő funkció a prediktív szövegbevitel, vagyis, hogy egy olyan adatmező esetén, ahol csak bizonyos értékek szerepelhetnek, ha a felhasználó elkezdi begépelni az értéket, a rendszer képes automatikusan kiegészíteni azt. Ez nagyban fokozza a hatékonyságot és egyúttal megelőzi a hibás adatbevitelt. Emellett a rendszer képes a feltételes kötelezőséget is kezelni, azaz, hogy egy adatkategória csak akkor válik kötelezővé (pl. forrás), ha egy másik adatkategória jelen van (pl. kép).

Hasonlóan a végfelhasználót segítő szövegek szerkesztéséhez, az adatmezők kitöltését egy HTML szerkesztő támogatja, így lehetőség van karakterszintű (pl. félkövér, dőlt vagy aláhúzott szedés), illetve bekezdésszintű formázást végezni (pl. számozott listaelemek használata). Ez nagyban hozzájárul ahhoz, hogy a végfelhasználót segíteni lehessen az adatmezőben lévő információ befogadásában a szöveg strukturálásával és ezáltal a figyelem megfelelő irányításával.

A *quickTerm* lehetőséget biztosít a tömeges szerkesztési műveletekre. Ennek beállítását egészen pontosan meg lehet határozni, különböző adatkategóriákra lebontva. A keresés és csere opció nem csak a teljes szavas egyezésekre terjed ki, hanem reguláris kifejezések használatával szöveges mintázatokat is könnyen meg lehet határozni. A szöveges módosításon túl pedig a bejegyzés, a nyelvi indexálás, illetve a terminus szintjén is megoldható az egyszerre történő átkategorizálás.

A tömeges műveletekhez szorosan kapcsolódik a minőségbiztosítási funkciók megléte is. Ez egyrészt megmutatkozik mikroszinten, vagyis lehet validációs mechanizmusokat beállítani az egyes adatmezőkre is (pl. milyen értéket lehet érvénytelennek tekinteni, illetve miként jelezze a rendszer, ha valamelyik kötelező adatkategória üresen marad). Az ezekhez kapcsolódó hibaüzenetek jórésze is testreszabható, így a szerkesztő felhasználók számára

is világossá válhat, hogy milyen módosításra van szükség. A tájékozódást vizuálisan segíti például a kötelezően kitöltendő adatmezők más színnel való megjelölésének lehetősége is. A minőségbiztosítás továbbá makroszinten is megmutatkozik, ugyanis lehetséges konzisztenciaellenőrzéseket végezni a bejegyzésen belül, illetve azok között is, valamint a bejegyzéseket összekötő vagy az adatbázison kívülre mutató hiperhivatkozások működképességét is lehet ellenőrizni. A rendszer további nagy előnye, hogy ezeknek az automatikus ellenőrzési műveleteknek a lefuttatását lehet ütemezni és az eredményt e-mailes formában el lehet juttatni a megfelelő felhasználói csoport(ok)hoz (ez leginkább a terminológiai projektvezetőket jelenti a gyakorlatban).

Az automatizálás másik fontos vetülete a munkafolyamatok kialakítása és kezelése, amelyre a *quickTerm* számos, igen részletesen finomhangolható megoldást nyújt mind a bejegyzések, mind pedig a fogalmi hálók vonatkozásában. Az elérhető munkafolyamat-típusok alapvetően két nagy csoportra oszthatók: a kívülről, a végfelhasználó oldaláról érkező terminushozzáadással kapcsolatos kérelmek, illetve a már meglévő bejegyzések célnyelvi megfeleltetésével kapcsolatos feladatok. Előbbi esetében a rendszer automatikusan ellenőrzi, hogy a beküldendő terminus már létezik-e a rendszerben, és pozitív válasz esetén erre figyelmezteti a felhasználót. A másik csoport pedig a terminológiai munkafolyamat lépéseit képes technológiailag megvalósítani, vagyis olyan lépések határozhatók meg, ahol a folyamat különböző szereplőinek más-más feladatuk van, ezeket az érintett felhasználók pontosan nyomon követhetik, valamint a szoftver lehetőséget biztosít az egyes lépéseknél a munkafolyamat aktuális állapotára utaló adatkategóriák automatikus hozzárendelésére az adott bejegyzéshez.

Valamennyi munkafolyamat-típust be lehet állítani akár nyelvenként vagy doménenként eltérő módon. A terminológiai munkafolyamatok egyik gyakori kezdeti lépése a külső forrásból való importálás; a rendszer itt is számos beállítási lehetőséget kínál kezdve az előzetes duplikátumszűréstől egészen az egyéni munkafolyamatbeli kategóriák hozzárendeléséig. Ennek nyomán elkerülhető, hogy egyszerre nagy mennyiségű adat kerüljön be ellenőrizetlenül az adatbázisba, amennyiben az előre meghatározott feladatláncnak maradéktalanul teljesülnie kell, hogy a folyamat végén az importált terminusok immár kidolgozott bejegyzések formájában kerüljenek publikálásra a végfelhasználók felé. Az egyes szerkesztői felhasználók munkáját pedig szintén megkönnyíti a rendszer oly módon, hogy lehetőséget biztosít számukra, hogy az új feladatról vagy munkafolyamatbeli lépésről e-mailes értesítést kapjanak, és a rendszerbe belépve pontosan átlássák, hogy milyen feladaton és milyen határidővel kell dolgozniuk.

Mindezek nyomon követésére és a terminológiai projekt előrehaladásának vizsgálatára a rendszer számos statisztikai funkciót is kínál. Ezek nem pusztán a szerkesztési munkát képesek megjeleníteni különböző szempontok mentén, hanem lehetőséget biztosítanak a végfelhasználói viselkedés mérésére is, amely reprezentatív képet adhat a terminológiai projektben résztvevő valamennyi csoport számára.

Az általánosabb technológiai szempontok közül kiemelendő a vizuális testreszabhatóság. Ez nem pusztán azt jelenti, hogy az adott szervezet használhatja a saját logóját, hanem az egész adatbázis színvilágát is megváltoztathatja a kezdőképernyőtől kezdve egészen a legapróbb adatmező betűtípusáig. Ennek révén – különösen egy nemzeti adatbázis esetén – még nagyobb hangsúlyt lehet fektetni az egyediségre, valamint a felhasználók igényeihez való alkalmazkodásra.

Ez a fajta rugalmasság megjelenik az adatok kinyerésekor, illetve a terminológiai keretrendszer más rendszerekkel való összekapcsolásának kérdésében is. Az exportálást és rendszerközi szinkronizálást különálló modulok teszik lehetővé (Publishing és Excelling), amelyek egyrészt a más rendszerekben is megszokott fájlformátumok (pl. TBX, Excel fájl vagy egyénileg testreszabható XML fájl), vagy speciálisan két rendszer összekapcsolására létrehozott alkalmazásprogramozási interfész (API) segítségével történik. Az adatkinyeréssel kapcsolatos beállítások sokszínűségét az is szemlélteti, hogy nagyon pontosan meg lehet határozni, milyen keresési és/vagy szűrési kategóriák mentén szeretnénk ezeket a műveleteket elvégezni vagy akár az előző kinyerési feladathoz képest csak az új adatokra akarjuk korlátozni a műveletet, valamint az exportálás, illetve az adatcsere nem pusztán a keretrendszer adatbázis részére vonatkozik, hanem a fogalmi hálókra is, amelyeket a rendszer az ontológiai és tudásmenedzsment-rendszerekben használatos fájlformátumok (pl. RDF) formájában valósít meg.

Az itt ismertetett funkciók természetesen nem kimerítőek, de már ennyiből is látható, hogy a *quickTerm* rendszere képes egy összetett alapadatbázist és komplex, többszereplős folyamatokat kezelni, mindezt kiegészítve olyan lehetőségekkel, amelyekkel az adatbázis nem csak a humán, de a gépi felhasználást is képes támogatni, így biztosítva azt, hogy egy nemzeti terminológiai adatbázis még hosszú éveken át releváns maradjon a felhasználási területek széles spektrumán. Hátrányként elsősorban a költségvetési szempontok merülhetnek fel, továbbá fontos kiemelni, hogy komoly idő- és erőforrás-ráfordítást igényel a rendszerben rejlő széleskörű lehetőségek minden részének megismerése és finomhangolása.

4. Következtetések

A tanulmányban elemzett két terminológiakezelő rendszerrel kapcsolatos megállapítások két fontos szempontra világítottak rá.

Egyrészt már elérhetőek olyan szoftverek, amelyek alkalmasak a végfelhasználó és a szerkesztő felhasználó igényeinek való megfelelésre, sokszor meglehetősen modern és jól működő funkciókat felmutatva, ezért nem érdemes a nulláról indulni és saját rendszer fejlesztésébe kezdeni, hiszen az rengeteg erőforrást igényelne és hosszadalmas munkát foglalna magába.

Másrészt itt is megmutatkoznak a nyílt forráskódú és a kereskedelmi szoftverek közötti alapvető különbségek: bizonyos szintig mindkét szoftver képes ugyanazokat a funkciókat biztosítani (még ha más technikai megoldások mentén is), ugyanakkor a vizsgálatból szintén látható, hogy az ingyenes szoftvernél hamar korlátokba ütközünk és kompromisszumokat kell kötni, míg az elemzett kereskedelmi szoftver esetén több és többféle lehetőség áll rendelkezésünkre a terminológiai projekt megvalósítására – igaz, ennek nagyobb is a költségigénye.

Mivel egy nemzeti terminológiai adatbázis felhasználói köre igen széles, mind a felhasználói, mind a terminológusi oldalon, ezért érdemes a lehető legrugalmasabb és legszéleskörűbb funkciókat nyújtó rendszert választani, hiszen a jövőbeni folyamatos fejlesztések reményében így lehet biztosítani, hogy a nemzeti terminológiai adatbázis ki tudja használni az éppen aktuális (nyelv)technológia lehetőségeit.