

A WIPO Pearl szabadalmi terminológiai adatbázis bemutatása

Tamás Dóra Mária

E-mail: tamas.dora.maria@gmail.com, tamas.dora@offi.hu

Kivonat: A tanulmány célja a Szellemi Tulajdonjog Világszervezet (World Intellectual Property Organization, WIPO) szabadalmi fordítási osztálya által (Patent Cooperation Treaty Translation Division)^{1,2} megvalósított, nyilvánosan elérhető WIPO Pearl szabadalmi terminológiai adatbázis bemutatása. A terminológiai adatbázis leírása és elemzése a Tamás és Sermann (2019a, 2019b) kísérleti szempontrendszer alapján készült, amelynek célja, hogy a nyomtatott és elektronikus szótárak elemzési módszereit követve a nagy szervezetek online terminológiai adatbázisok elemzése is lehetőség szerint szakmailag objektív szempontok alapján készüljenek. Ennek megfelelően az írás összegzi a terminológiai adatbázissal kapcsolatos háttérinformációkat, a technikai paramétereket, az adatbázis tartalmáról és használatáról szóló információkat, továbbá kitér a szabadalmi terminológia sajátos jellemzőire.

Kulcsszavak: összetett terminológiai adatbázis, fogalmi háló, nyelvi és fogalmi keresés, szabadalom, szabadalmi terminológia

1. Bevezetés

A WIPO Pearl szabadalmi terminológiai adatbázis jellemzőit a Tamás és Sermann (2019a, 2019b) által kidolgozott szempontrendszer alapján mutatom be. A WIPO Pearl adatbázis bemutatása és elemzése több szempontból is érdeklődésre tarthat számot. Egyfelől összetett terminológiai eszközzel van szó, amely a hagyományos nyelvi keresési módokon felül a fogalmakra való keresést ontológiai eszközökkel is támogatja. Másfelől a szabványokban és a bevált gyakorlatokban közzétett terminológiai szerkesztési szempontokat gyakran egyedivé teszik egy adott domén terminusainak és fogalmainak sajátosságai. Ez igaz a szabadalmak világában is, amely terület sajátos jellemzői bizonyos mértékben eltérő feldolgozási és szerkesztési módszereket kívánnak meg, így e sajátosságokat követve alakították ki a terminológiai adatbázis felépítését és tartalmát.

Hivatkozás: Tamás D. 2021. A WIPO Pearl szabadalmi terminológiai adatbázis bemutatása. *Fordítástudomány* 23. évf. 1. szám. 49–62.

DOI: <https://doi.org/10.35924/fordtud.23.1.3>

1.1. Az elemzési szempontok mögött meghúzódó megfontolások

A leíráshoz és elemzéshez alkalmazott kísérleti szempontrendszer elsősorban nagyobb szervezetek központi, nyilvános elérhetőségű terminológiai adatbázisainak bemutatására szolgál, és a nyomtatott és elektronikus szótárak kritikai szempontrendszeréből kiindulva született abból a célból, hogy a terminológiai adatbázisok megítélése sem szubjektív, hanem lehetőség szerint objektív szakmai szempontok alapján történjen (Fóris és Rihmer 2007, Gaál 2012).

Központi adatbázisokról lévén szó, a különböző elemzési kategóriák az online felületre fókuszálnak, hiszen ahogyan azt megfigyeltük, az egyes terminográfiai bejegyzések belső elrendezése (mikrostruktúrája) az online felületen általában közvetlenül nem látható, és így nem vizsgálható. Gondoljunk például arra, hogy a hierarchikus elrendezésű, három szintű szerkesztési struktúrában belül a definíció és a terminus adatmezők akár két különböző szerkesztési szinten is szerepelhetnek. Ezek elhelyezkedése akkor írható le, ha rendelkezésre áll valamilyen egyéb, a nem látható jellemzőket bemutató leírás a belső szerkezetről. A WIPO Pearl esetében szerencsés a helyzet, mert nemcsak az egyes adatmező-kategóriákat leíró felhasználóknak szánt használati útmutatóhoz (<https://www.wipo.int/reference/en/wipopearl/guide.html>) férhetünk hozzá, hanem tanulmány formájában az adatbázis belső struktúrájának leírásához is, amelyet a szerzők éppen azért tettek közzé, mert a fent leírtakkal tisztában voltak (Valentini et al. 2016). Az említett tanulmány szerzői, akik egyben a belső, szerkeszthető PCT terminológiai adatbázis és az online elérhető, kizárólag adatközlő felületként működő WIPO Pearl fejlesztői is, írásukban a belső adatbázis létrejöttének és fejlesztésének körülményein kívül ismertetik azt, hogy a hármas hierarchikus struktúrában (bejegyzés, nyelvi indexálás és terminus szintje) szerkesztési szintenként milyen kötelező vagy fakultatív adatmezőtípusok szerepelnek, és mennyiben követték az adatbázis szerkezetének kidolgozásakor az ISO kategóriákat, továbbá bemutatják a szabadalmi fordítások területén alkalmazott terminológiai feldolgozás sajátosságait.

Bár a jelen tanulmány elsősorban a WIPO Pearl szabadalmi terminológiai adatbázis online felület fő jellemzőinek leírására fókuszál, a fent említett tanulmányban szereplő háttérinformációk (Valentini et al. 2016), továbbá a WIPO *PCT Translation Division* elnevezésű szabadalmi fordítási osztály munkatársaival folytatott egyetemi együttműködés és konzultáció (Tamás 2020) lehetővé tette, hogy igény szerint bekukkantsunk a színtfalak mögé.

2. A WIPO Pearl terminológiai szabadalmi adatbázis részletes elemzése

Érdekes elsőként definiálni, hogy mit is értünk terminológiai adatbázison: „olyan elektronikusan tárolt terminológiai adatok összessége, amely az onomasziológiai szemléletet követve készült, megalkotásának előfeltétele az adott szakterület fo-

galmi rendszerének felvázolása. Egy vagy több szakterület terminusait és a hozzájuk tartozó definíciókat tartalmazza egy vagy több nyelven.” (Sermann és Tamás 2010: 113; Tamás és Sermann 2019a: 33).

2.1. A WIPO Pearl terminológiai adatbázissal kapcsolatos háttér-információk

A tágabb háttérét tekintve a WIPO Pearl terminológiai adatbázis a szabadalmi terminusok területén egyedülálló terminológiai segédeszköz, amely elsősorban a szabadalmakban szereplő fogalmak és terminusok közzétételére szolgál. A szabadalmak szövege terminusokban gazdag szövegtípus, a 2000-es évek elején születtek meg a WIPO részéről az első erőfeszítések a szabadalmi terminológia szoftverrel támogatott kétnyelvű kivonatolására a fordítási folyamat elősegítése céljából, és született meg az első, belső adatbázis-struktúra (lásd *PCT Termbase*). 2005-től már megvalósult az első változat belső használata, de a szisztematikus fejlesztése csak 2007-ben kezdődött meg. Eleinte önálló terminológiai adatbázisként működött, később vált egy terminusfelismerő funkcióval rendelkező CAT-rendszer részévé. Első pozitív hatásai a konzisztens terminushasználatban és időmegtakarításban jelentkeztek. 2010-ben született stratégiai döntés az online változat létrehozásáról és a terminológiai adatok publikussá tételéről, ami 2014-re valósult meg (Valentini et al. 2016). A terminológiai irányzatokat figyelembe véve elsősorban fordítói, illetve fordításorientált terminológiai adatbázis. Bár belső változata elsősorban a fordítási munkafolyamatot hivatott elősegíteni, a nyilvánosan elérhető felület működése révén nemcsak a szabadalmi terminológia fejlődését támogatja, hanem indirekt módon a fejlesztők technológiai és az egyes nyelvek szaknyelvi és terminológiai fejlődését is, tehát akár nyelvpolitikai célokat is szolgál.

Szűkebb háttérét tekintve érdemes megvizsgálni a terminológiai adatbázis-előállító jellemzőit. A WIPO Pearl fenntartója a Szellemi Tulajdon Világszervezete, amelyet 1967-ben alapítottak az ún. Stockholmi egyezmény³ alapján, hogy világszerte védje a szellemi tulajdont, és ösztönözze, elősegítse a kreatív tevékenységeket. Alapítása után hét évvel, 1974-ben lett az ENSZ szakosított szervezete, székhelye Genfben található, tagországainak száma 193 (<https://www.wipo.int/about-wipo/en/>).

A WIPO szervezetén belül a szabadalmi fordítási osztály (*PCT Translation Division*) felel a Szabadalmi Együtműködési Szerződés keretein belül benyújtott nemzetközi szabadalmi kérelmek fordításáért. A WIPO Pearl fejlesztője ez utóbbi fordítói osztály, amelynek fő tevékenységi körébe a szabadalmak címének, kivonatának és mellékletének publikálás előtti angol és francia nyelvre történő fordítása, továbbá szabadalmazhatósági véleményezések fordítása tartozik. Hatalmas volumenű fordítási mennyiségről van szó, 2018-ben 610.000 darab fordítás készült el, ami 160 millió lefordított szónak felel meg. A fordítási osztály nevében a PCT rövidítés arra utal, hogy tevékenységét az 1978-ban hatályba lépett Szabadalmi Együtműködési Szerződés (*Patent Cooperation Treaty*) hatálya alatt kezdte meg

és az abban megállapított nyelvek szerint végzi, amelynek idén ünnepelték ötvenedik aláírását. Ez a nemzetközi szerződés lehetővé tette a szabadalmak elnyerésének egyszerűbbé, gyorsabbá, olcsóbbá és megbízhatóbbá tételét abban az esetben, ha egy találmányt több országban kívánnak szabadalmaztatni. 152 ország tagja, Magyarország 1980 óta aláírója e szerződésnek. E szerződés képezi alapját a nemzetközi szabadalmi bejelentésnek (*international patent application*), amelynek benyújtása általában a nemzeti hivatalnál történik meg (lásd Szellemi Tulajdon Nemzeti Hivatala, SZTNH, <https://www.sztnh.gov.hu/hu>). A szabadalmi kérelem bejelentését a kérelmezők elsőként a nemzeti (regionális) szabadalmi hivatalba nyújtják be, és ezt követően a kérelmezőnek 12 hónap áll rendelkezésre, hogy eldöntse, hogy a kérelmét nemzetközi szinten is érvényesíteni kívánja avagy sem a Szabadalmi Együttműködés Szerződés értelmében és ezzel automatikusan 152 államban elindul az eljárás. Ilyen esetben minden államban az első, nemzeti hivatalnál tett benyújtási időpontot veszik figyelembe. Ennek részeként nemzetközi újdonságkutatás és nemzetközi közzététel történik, illetve a bejelentő kérhet ún. nemzetközi elővizsgálatot <https://www.sztnh.gov.hu/hu/patent-cooperation-treaty-szabadalmi-egyuttmukodesi-szerzodes>. Fő kérelmezői közé tartoznak Kína, az Egyesült Államok, Japán, Németország, Korea, Franciaország (<http://www.wipo.int/ipstats/en/>), a szabadalmak adatai a PATENTSCOPE szabadalmi tárban férhetők hozzá (<https://www.wipo.int/PATENTSCOPE/en/>⁴).

A szabadalom mint műfaj megértése összefügg a találmányok világával. „A szabadalom a találmányra engedélyezett kizárólagos jog, amelynek alapján a jogosult a találmányt hasznosíthatja, és mindenki más el van tiltva annak hasznosításától” (Gács 2011: 5). A találmány (*invention*) a termékre vagy eljárásra vonatkozó új műszaki megoldás, és az azon fennálló jogi oltalmat nevezik szabadalomnak (*patent*) vagy szabadalmi oltalomnak. A szabadalom a bejelentés közzétételével keletkezik. Egy szabadalom bizonyos feltételek fennállása mellett szabadalmazható: új (nem hozzáférhető), feltalálói tevékenységen alapul (nem nyilvánvaló) és iparilag alkalmazható.

A szabadalmi bejelentés (*patent application*) részei általában az alábbiak:

- bejelentési kérelem (*demand*; a magyarországi nemzeti kérelemnek részét képezi);
- leíró rész igénypontokkal – többlethatás leírása (*description with claims*);
- kivonat (*abstract*);
- rajz és egyéb mellékletek (*drawings*). (lásd bővebben: <https://www.wipo.int/pct/en/texts/glossary.html>; Gács 2011).

A fentieknek megfelelően a WIPO Pearl terminológiai adatbázis elsődleges célcsoportját a szabadalmak fordítói és az adott szakmai képviselői, például jogászok, mérnökök, dokumentumszerkesztők jelentik. Felhasználási célját tekintve a belső felhasználók, fordítók számára lehet előíró jellegű, a külső felhasználók számára leíró jellegű, használata ajánlott.

2.2. A terminológiai adatbázis technikai paraméterei

Az adatbázis technikai paramétereit tekintve az online felület adatai egy saját, egyedi fejlesztésű, szerzői joggal védett szoftver révén férhetők hozzá. A WIPO Pearl nyilvános, lekérdezésre alkalmas, web-alapú felülete mellett létezik egy fordítástámogató szoftverrel használható belső adatbázis. A WIPO Pearl online lekérdezhető felülete használható önmagában, továbbá a belső, PCT Termbase elnevezésű adatbázis a szabadalmi fordítók számára egy komplex nyelvtechnológiai jellegű háttértámogatásnak köszönhetően egy integrált szoftverfelületen használható. A WIPO Pearl dizájnya utoljára 2019-ben frissült, adatai folyamatosan frissülnek, az utolsó adatfeltöltés ideje 2020. szeptember. Adatai további futtatási platformokon is olvashatók, mint okostelefon vagy táblagép (https://www.wipo.int/reference/en/wipopearl/news/2019/news_0003.html). A szervezetbe bevezetett terminológiai és fordítástámogató segédeszközök alkalmazásáról részletesen ír Caffrey és Valentini (2019).

2.3. A terminológiai adatbázis tartalmáról szóló információk

A WIPO Pearl tartalmának fő jellemzőit tekintve fogalomközpontú elvek szerint szerkesztett terminológiai adatbázis. Felépítését, azaz a szerkezeti egységek számát tekintve összetett terminológiai adatbázis, amely a terminológiai bejegyzéseket tartalmazó klasszikus terminológiai adatbázis mellett fogalmi hálókat is tárol. Ennek megfelelően a nyelvi keresési felület (*Linguistic Search*) mellett a fogalmi hálókból való keresést (*Concept Map Search*) is felkínálja. Ez utóbbi keresési felület domén, aldomén és nyelvek szerint mutatja be a fogalmi hálókból a terminusok közti kapcsolódási pontokat. Ezen relációk típusai két nagy alcsoportra bonthatók:

- hierarchikus viszonyok ábrázolására, és ezen belül alá- és fölérendeltség vagy rész-egész viszonyok ismertetésére, és
- asszociatív (azaz tematikus, nem hierarchikus viszonyt ábrázoló) relációkra.

Az egyes terminográfiai bejegyzések felépítése részletesen kidolgozott, megfelel az elementaritás (egy adatmezőben egy adat szerepeltetése), a granularitás (az adatok kidolgozottságának mértéke legyen megfelelő a túl sok alkategória elkerülése érdekében) (lásd Reinke 2012: 102), és számos adatmező-kategóriát tartalmaz. Ha megvizsgáljuk a terminológiai adatbázis szerkezetének részletes jellemzőit, akkor megállapíthatjuk, hogy a szótáraknak megfelelő megastruktúra szerkezeti szintjén a nyitó honlapon egy rövid leírást találunk az adatbázis tartalmáról és céljairól, továbbá egy link segítségével átléphetünk a használati útmutatóba (<https://www.wipo.int/reference/en/wipopearl/guide.html>). A nyitó honlap a WIPO szervezet által fenntartott honlapok részét képezi.

A WIPO Pearl terminológiai adatbázisban tárolt nyelvek száma tíz, és meg egyezik az ENSZ hivatalosan használt nyelveivel (angol, francia, spanyol, arab, kínai, orosz) és a Szabadalmi Együttműködési Szerződés szerint működő szabadalmi fordítói osztály (*PCT Translation Division*) munkanyelveivel (német, portugál, japán, koreai). Az adatbázis 29 doménbe és 311 aldoménbe sorolva tárol terminusokat, ezek közül a fő tárgykörök (domének) a műszaki és természettudományi tárgykörök, illetve a szellemi tulajdon mint jogi domén. Az online felületen a validált (ellenőrzött és jóváhagyott) terminusok bejegyzései tanulmányozhatók. A fogalmi keresés felületén könnyen megfigyelhető, hogy bizonyos tárgykörök a technológiai fejlődés és a szabadalmi igények függvényében jobban kifejtettek, míg más területek kevesebb terminust tartalmaznak, idővel pedig új területek jelenhetnek meg. A doménrendszer kialakításának alapját az 1971-ben a Strasbourgi Megállapodással megalapított IPC-rendszer mellett (*International Patent Classification*, <https://www.wipo.int/classifications/ipc/en/>; <https://www.sztnh.gov.hu/hu/nemzetkozi-szabadalmi-osztalyozas-nszo-international-patent-classification-ipc>), az ISO áru- és szolgáltatási osztályozása, illetve egyéb taxonómiák képezik (Valentini et al. 2016).

A WIPO Pearl jelenleg 190 000 terminust és 24 000 fogalmat (2020. évi szeptember havi adat) tartalmaz, terminusai elsősorban szabadalmakból származnak. Makrostruktúrája alapján a nyelvi keresés (*Linguistic Search*) egyszerű és összetett keresési lehetőséget is felkínál. Ez utóbbi esetben lehetővé teszi a forrás- és célnyelvi, a domén szerinti, az adott terminusra történő szűkített keresést, a rövidítések keresését és néhány esetben a gépi fordítási találatok megjelenítését is. A keresőmezőbe beírt terminusra adott találatokat egyszerűen áttekinthető, felsorolás jellegű találati listában kínálja fel. A felső sorban az átláthatóság érdekében külön feltünteteti egy sorban a több oldalon felsorolt terminusokat. A fogalmi hálókban való keresésnél (*Concept Map Search*) beállíthatjuk a kívánt nyelvet, és a tárgykörök rövidítésével ellátott színes körökre kattintva kiválaszthatjuk a megnyitandó domént, aldomént és fogalmi kört, amely utóbbi már hálózatos formában jelenik meg. A terminusra egérrel ráállva láthatjuk a terminus szinonimáját, rákereshetünk továbbá arra, hogy milyen terminusok érintésével juthatunk el egyik terminustól a másikig, ami a szabadalmi tárban való kulcsszavas kereséshez is segítséget nyújthat. Pozitívumként megemlítendő, hogy a két keresési felület között az átjárás biztosított.

Az egyes szótári szócikkek, így a mikrostruktúra szintjének megfelelő terminográfiái bejegyzések kétféle módon jelenítik meg az adatokat: egy egyszerűsített és egy részletes formában (*full record display*). Jelen tanulmányban elsősorban az online felület adatmezőtípusaira koncentrálnak, a többi adatmezőtípus leírása, illetve az, hogy mennyiben felelnek meg a korábbi ISO szabványokban ajánlott kategóriáknak, az Valentini et al. (2016) írásában tanulmányozható.

A találati listán belül szereplő egyszerűsített nézet adatmezőtípusai a terminus, a megbízhatóság (például 3/4), más adatállományokba átvezető linkek, a definíció, a kontextus és a forrás. Ennek kattintással követő kibontását követően

a részletesen megjelenített bejegyzések adatmező kategóriái a következők: domén, aldomén, eredeti rögzítési nyelv, nyelv, a fogalmi hálóra átvezető link, terminus, terminus típusa (fő terminus, szinonima), terminus leírása (rövidítés), használati címke (megengedett, archaikus, terminusjavaslat a fordítói osztály részéről, ajánlott, szabványosított használat), terminus megbízhatósága négyes skálán, utolsó módosítás dátuma, kontextus, kontextus forrása és a terminushoz fűzött megjegyzés.

Az adatbázisban található terminusok szófajukat tekintve elsősorban főnevek, de lehetnek igék, melléknevek, határozószók vagy több tagból álló lexémák, kollokációk is (Wright és Budin 1997: 3). A WIPO munkatársainak a gyakorlatban a terminus elsősorban főnév, ritkán ige vagy nagyon ritka esetben önmagában álló melléknév (Valentini et al. 2016).

Ezen a ponton érdemes kitérnünk röviden a szabadalmi terminusok és a szabadalom mint műfaj jellemzőire. A szabadalom egy komplex szövegtípus, határterületet képez a tudományos, a műszaki és a jogi szaknyelv között (Valentini et al. 2016). Megalkotását befolyásolják a nemzeti jogszabályok is, és különböző részekből áll, amelyek közül a leíró rész (*description*) inkább tudományos-műszaki, míg az igénypontokra (*claims*) inkább a jogi szakszöveg tulajdonságai jellemzőek.

A szabadalmak speciális nyelvezete onnan ered (amelynek angol megnevezése az ún. „patentese”), hogy egyfelől szükséges a találmányok közzététele, bemutatása oly módon, hogy azt egy szakember a leírás alapján megértse, tehát a jogszabályi előírásoknak megfelelően beazonosíthatónak kell lennie, ami pontos terminushasználatot követel meg, másfelől védelmet kell nyújtania az adott területen a versenytársak hasonló találmányaival szemben, ami egy tágabban értelmezhető, leíró jellegű terminushasználatot eredményez. Ez utóbbi nyelvhasználat megnehezíti a terminusok kivonatolását, és a fő terminusok és a használatra ajánlott terminusok kiválasztását a számos változat közül (Valentini et al. 2016). Ez a fajta megközelítés nemcsak az angol nyelv esetében igaz, hanem a szabadalmak funkciója miatt a többi nyelv esetében, így a magyar szabadalmak esetében is.

A terminusok validálásakor fontos szempont az előfordulási gyakoriságon túl, hogy a kiválasztott terminus ne legyen túlzottan specifikus, ne csak egy találmányhoz kötődjön, ugyanakkor ne legyen túl általános sem, hanem konkrét, beazonosítható fogalmat jelöljön. A túlságosan tág értelmezésű, leíráshoz használt szabadalmi terminusok általában több tagból álló általános leíró összetételek és a következő tagokat tartalmazzák: szó + „,means’, ’system’, ’device’” (Valentini et al. 2016: 184), de ide sorolhatók még további tagok is (lásd *module, assembly, device, unit, portion, set, block, reservoir* stb.). A különböző tudományos és műszaki szakszövegek tulajdonságairól ír Olohan (2016), aki könyvében külön fejezetben értekezik a szabadalmak fordításáról, annak jellemzőiről és minőségük jelentőségéről.

Visszatérve az adatmező-típusokra, az adatbázis megkülönböztet kötelező és fakultatív adatmező-kategóriákat, amelyek közül a kötelező adatmezőtípusok a következők (Valentini et al. 2016: 197):

- a bejegyzés szintjén (*entry level*): bejegyzési osztály (validálás státuszai: új, validált, félig-validált és szerzői joggal védett), domén, aldomén, eredeti rögzítés forrásnyelve;
- a terminus szintje (*term level*): terminus, terminus státusz, használati címke, terminus típusa, forrás.

Az adatmezők besorolása megfelel a hármas hierarchikus struktúrának, amelyen belül a nyelvi indexálás szintjén (*language/index level*) nem találunk kötelező adatmezőtípusokat. Az utóbbi szerkesztési szintre besorolt fakultatív jellegű, a belső adatbázisban szereplő „*Transfer Comment*” (átváltási megjegyzés) tartalmaz ugyan a fogalmi egyenértékűségre, azaz az ekvivalenciára vonatkozó információkat, de szintén fakultatív a definíció mint adatkategória (Valentini et al 2016: 197). Ez szokatlan megoldás a fogalomközpontú terminológiai adatbázisokban, hiszen a fogalmat leíró definíció hagyományosan a terminológiai adatbázisok központi eleme. A definíció szerepét ugyanis itt pragmatikus okokból átvette egyfajta kompromisszumként egy speciális kategória, az ún. definíciós kontextus, azaz a fogalmi jellemzőkben gazdag kontextus. Ennek oka abban keresendő, hogy a definíciók kidolgozása jelentős munkaigényű, ezért idő- és pénzigényes, szakértői hálózat fenntartását igényli nyelvenként és doménekenként, a gyakorlatban azonban a technológia és az ipar fejlődése gyors, a fordítások pedig szoros határidővel készülnek (Valentini et al. 2016).

Nézzük meg, hogy mit értünk pontosan definíciós kontextus alatt (*defining context*, lásd 2.3. pont alatt). Hagyományosan a kontextus igazolja, hogy a terminus használatban van, és bemutatja a terminus használatát közvetlen szöveggörnyezetében, azaz hogyan illeszkedik a terminus az adott mondatba. Drewer és Schmitz (2017: 67) a „*Definitionskontext*” változatot nem nevezi hagyományos értelemben definíciónak, és általában véve olyan mondatot értenek alatta, amely megnevezi az adott terminust, információt szolgáltat ugyan a terminus által jelölt fogalomról, de nem alkalmas a fogalom pontos behatárolására. Tehát tartalmaz definícióra jellemző elemeket, de nem feltétlenül meríti ki egy definíció teljességre törekvését a fogalom teljes körű behatárolásához, és nem is egyezhet meg egy szakember által alkotott, a fogalom pontos és teljes behatárolására törekvő definícióval. Valentini et al. (2016: 189) ezeket a kontextusokat ismeretekben gazdag kontextusnak („*knowledge-rich contexts*”) is nevezi, és ennek használata mellett érvel abból a szempontból, hogy használatuk kevésbé idő- és költségigényes, és lefedettséget biztosít.

A szabadalmi kérelmek leíró része olyan információkat tartalmaz, mint például a találmány tárgya és alkalmazási területe, a technika állásának ismertetése (milyen hiányosságokat pótol a találmány), a megoldások ismertetése fő- és aligénypontokkal, az ábrák felsorolása, az oltalom terjedelmét alátámasztó példák és az előnyös hatások bemutatását. Ez a szabadalmi kérelmek legalkalmasabb része a terminusok és definíciós kontextusok kivonatolására. Éppen ezért a szabadalmakon belül az eredeti, anyanyelvi szerzőtől származó leíró rész számít elsőd-

leges forrásnak, hiszen ez a rész tartalmazza legnagyobb valószínűség szerint a felhasználható terminusokat, definíciós kontextusokat és definíciókat. Másodlagos források közé tartoznak a tanulmányok, szakdolgozatok, doktori értekezések és tankönyvek, amelyek közül számos elérhető például a Google Books, Google Scholar és Google Patents felületén.

A fogalmi keresés (*Concept Map Search*) alatt kínált találatokban a terminusok által jelölt fogalmak közti relációkra láthatunk rá. A fogalmi háló jelentősége abban mutatkozik meg, hogy a fogalmak nem önmagukban, hanem mindig egymáshoz kapcsolódva léteznek (lásd ISO 704:2009:8)⁵. A terminológusok az egy-egy doménbe és aldoménekbe tartozó fogalmakat és terminusokat a fogalmi és terminológiai rendszerek feltérképezésével és kidolgozásával választják ki. A terminusok egymástól való elhatárolását és egymáshoz kapcsolódását már Wüster is táblázatok vagy ontológiai besorolások és elemek (gráfok/ágrajzok) formájában javasolta ábrázolni (lásd erről bővebben: Wüster 1985: 137–201). A rendszerben történő ábrázolás megkönnyíti a fogalmak behatárolását, a fölé-, alá- és mellérendeltségi viszonyok megállapítását. A klasszikus terminológiai munka tehát valójában a fogalmi háló felvázolásával kezdődik, ami segít az összes releváns terminus kiválasztásában, továbbá ellensúlyozza a terminusok különálló bejegyzésekben történő ábrázolását, ami töredezetté teszi az adatokra való rálátást, hiába jelöli egy külön adatmező-kategória a reláció típusát.

Az ISO 1087-1:2000 (*Terminology work — Vocabulary — Part 1: Theory and application*) az ontológiai elemek két fő ágaként az alábbiakat nevezi meg:

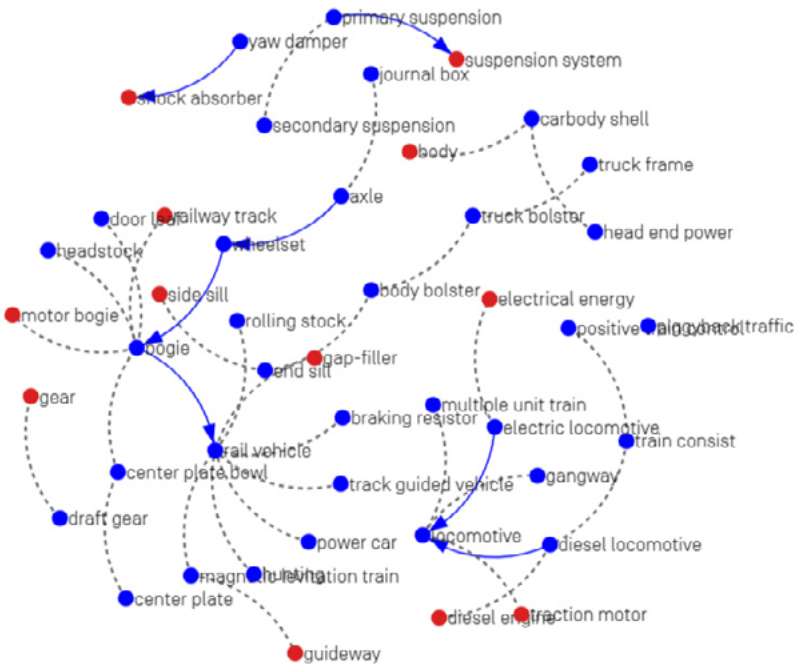
- hierarchikus (ezen belül alá- és fölérendeltség, rész-egész viszony) és
- asszociatív (például időbeli, ok-okozati összefüggések).

Ez a besorolás megegyezik a WIPO Pearl terminológiai adatbázisban a fogalmi hálóban szereplő relációtípusokkal, ahol harmadik kategóriaként szerepel még a másik doménhez tartozás jelölése (lásd *other subject field* 1. ábra). A fogalmi hálók ábrázolása a WIPO Pearl felületén egy doménen, egy aldoménen és egy nyelven belül történik. A WIPO munkatársai a relációkat elsősorban manuálisan rögzítik, és ez a tevékenység gyakran igényli tárgyi szakértők bevonását (Valentini et al. 2016). A fogalmak kb. 80%-a többnyire manuálisan feldolgozott, tehát kapcsolatban áll valamely másik fogalommal, további relációkat pedig mesterséges intelligencia (gépi tanulás) segítségével elemzik, ami megmutat további lehetséges kapcsolódási pontokat. Pozitívum, hogy mindig jelzi, ha a mesterséges intelligencia által generált találatról van szó. A relációk kisebb része tehát olyan mesterséges intelligencia alkalmazásának eredménye, amely saját fejlesztésű, és olyan fogalmi relációkat kínál fel, amelyeket még nem rögzítettek manuálisan. Ezeket „fogalmi felhőnek” (*concept clouds*) nevezik, és rendkívül innovatív és kísérleti jellegűek. Az ontológiai elemek azért igazán hasznosak, mert a modern számítógépes ontológiák fejlesztése és sikeres alkalmazásuk azon a felismerésen alapul, hogy a tudás megszerzése nem elegendő: „ismereteinket csak akkor alkalmazhatjuk a gyakorlatban, ha

ismerjük a különböző elemek közötti összefüggéseket és képesek vagyunk leírni azokat” (Fóris 2013: 422)⁶. Az ontológiai rendszerek online eszközeivel és ábrázolási módjaival foglalkozik még Ugrin (lásd Ugrin 2019: 99).

A fogalmi hálóban történő ábrázolás előnye a fordítások készítése szempontjából, hogy rálátást biztosít az adott fogalmi hálóra, mert bemutatja, hogy az adott doménen és aldoménen belül a terminusok és fogalmaik hogyan kapcsolódnak egymáshoz. Fordításnál ugyanis előfordul, hogy a másik nyelven csak egy gyűjtőfogalmat találunk meg elsőre, és így láthatjuk, hogy milyen terminusok tartoznak alá, és ez megkönnyíti a megfelelő terminus megtalálását. Ezen felül a fogalmi hálók nemcsak a fordítók számára hasznosak, hanem segítenek a szabadalmak kombinált kulcsszavas keresésekor (*WIPO Pearl Concept Path search*).

1. ábra: A vasúti jármű (*Rolling stock*) aldomén fogalmi hálója a WIPO Pearl felületén (<https://wipopearl.wipo.int/en/conceptmap>).



Ezzel együtt összességében elmondható, hogy az adatok kategorizálása betartja a mikrostruktúra szintjén a terminológiai adatmodellálás elveit: fogalomközpontúság, terminusautonómia (a fő terminus és szinonimák jelzése), hierarchikus szerkesztési mód, elementaritás (egy adatmezőben egy adat szerepeltetése) és a granularitás (a túl sok alkategória elkerülése egy adatmezőben) (Reinke 2012: 102).

A mezostruktúra, azaz az utalási rendszer szintjén már a találati lista felületén három link látható: a belső link átvezet bennünket a fogalmi hálók felületére, a két külső link a PATENTSCOPE szabadalmi tár és a Google képek keresési felületére.

2.4. A terminológiai adatbázis használatáról szóló információk

Az adatbázis felhasználóbarát jellegű, könnyen kereshető, egyértelműek a vizuális elemek. Az adatok folyamatosan bővülnek, a naprakészség a bejegyzéseken belül az utolsó módosítás dátuma elnevezésű adatmezőben ellenőrizhető.

Az adatok megbízhatóságát a pontos dokumentáció szolgálja, az egyes terminusok megbízhatóságát számokkal egy négyes skálán jelöli, amelyen belül az 1-es kategória alacsony megbízhatóságot, a források hiányát mutatja, a négyes kategória pedig a legmegbízhatóbb adatokat tünteti fel. Az adatbázisban szereplő definíciók ugyan ritkák és saját, belső forrásból szerkesztettek, az adatbázisban leginkább definíciós kontextusok szerepelnek. A kontextus megbízhatónak minősített forrásait szabadalmak, tudományos cikkek, szakdolgozatok, doktori tézisek, tankönyvek képezik (Valentini et al. 2019), amelyek visszakereshetősége a forrás adatmező segítségével megoldható. A megbízhatóság fontos kérdés, hiszen „egy célnyelven hibásan megadott terminus eldöntheti, hogy egy kérelmet a technika állása szerintinek vagy valóban újdonságnak tartanak-e?” (Valentini et al. 2016: 193). Egyetlen pontatlanság a rendkívüli értéket képviselő szellemi jogi termékek fordításában megváltoztathatja a szabadalom oltalmi körét (Ball 2013).

A WIPO Pearl adatbázis társadalmi haszna a fordítások, a szakemberek (adatbányászat és szabadalmak keresése a PATENTSCOPE szabadalmi tár linkjén keresztül) és terminológia-politikai célok mellett megmutatkozik még abból a szempontból, hogy fenntartója egyetemi és partnerségi együttműködésekre törekszik, egyetemi hallgatók számára csoportos bedolgozást és szakmai gyakorlatot tesz lehetővé (Tamás 2020).

3. Összegzés és következtetések

A Tamás és Sermann (2019a 2019b) elemzési és osztályozási szempontrendszer alapján bemutatott WIPO Pearl szabadalmi terminológiai adatbázis egy összetett, innovatív jellegű, plusz ismeretelemként ontológiai jellegű fogalmi hálókat tartalmazó terminológiai adatbázis. A tanulmány bemutatja a terminológiai adatbázissal kapcsolatos háttér-információkat, a technikai paramétereket, az adatbázis tartalmáról és használatáról szóló információkat. A szabadalmi terminusok sajátos jellemzői bizonyos mértékben egyedi szerkesztési módszereket kívánnak meg, az adatbázis felépítése és tartalma ennek kívánalmait követi. A WIPO Pearl adatbázis a fordítók mellett műszaki, jogi szakemberek számára is hasznos terminológiai eszköz, továbbá a szakszavak rendszerezésével és közzétételével terminológia-politikai célokat is szolgál.

Jegyzetek

- ¹ Szabadalmi Együttműködési Szerződés Fordítási osztálya, rövidített, angol nyelvű megnevezése: PCT Translation Division.
- ² Ezúton fejezem ki köszönetemet a WIPO munkatársainak, Geoffrey Westgate-nek (Head, Support Section, PCT Translation Division, WIPO) és Cristina Valentini-nek (Head, Terminology Unit, Support Section, PCT Translation Division, WIPO), hogy a cikk lektorálásával hozzájárultak e tanulmány megírásához.
- ³ A Szellemi Tulajdon Világszervezetét létrehozó 1967. július 14-i Stockholmi Egyezmény.
- ⁴ A Szellemi Tulajdon Nemzeti Hivatala felsorolja a nemzetközi és nemzeti szabadalmi adatbázisokat (<https://www.sztnh.gov.hu/hu/szabadalmi-adatbazisok>).
- ⁵ „Concepts do not exist as isolated units of knowledge but always in relation to each other”.
- ⁶ „The development of modern computerized ontology and its successful practical application in various fields is based on the recognition that obtaining knowledge is not enough; we can apply our knowledge in practice only if we know the connections among its various elements and are able to describe them”.
- ⁷ „A term misrepresented in a target language may be the difference between an application being prejudiced by prior art or being confirmed as indeed novel”.

Irodalom

- Ball, L. 2013. Is your patent quality lost in translation? *Intellectual Property magazine*. December 2012/January 2013. 54–55.
- Caffrey, C.–Valentini, C. 2019. Applications of technology in the Patent Cooperation Treaty (PCT) Translation Division of the World Intellectual Property Organization (WIPO). In: Minako O’Hagan (ed.) *The Routledge Handbook of Translation and Technology*. London: Routledge. 127–147. <https://doi.org/10.4324/9781315311258> (Letöltve: 2020. október 15.)
- Drewer, P., Schmitz, K.-D. 2017. *Terminologiemanagement Grundlagen - Methoden – Werkzeuge*. Berlin: Springer-Verlag.
- Fóris Á. 2013. Network Theory and Terminology. *Knowledge* Vol. 40. No. 6. 422–429.
- Fóris Á., Rihmer Z. 2007. A szótárak minősítési kritériumairól. *Fordítástudomány* IX. évf. 1. szám. 109–113.
- Gács J. (szerk.) 2011. *Útmutató feltalálóknak. A szabadalmi bejelentés*. Budapest: Szellemi Tulajdon Nemzeti Hivatala (https://www.sztnh.gov.hu/hu/kiadv/ingy_magy/szabadalmi_bejelentés_2011.pdf). (Letöltés ideje: 2020.09.01.)
- Gaál P. 2012. Szempontrendszer online szótárak minősítéséhez. *Magyar Terminológia* 5. évf. 2. szám. 225–250.
- Olohan, M. 2016. *Scientific and Technical Translation*. London/New York: Routledge.
- Reinke U. 2012: Terminologiearbeit für und mit Sprachtechnologie. In: Mayer F., Schmitz K.-D. (eds) *Terminologieprozesse und Terminologiewerkzeuge*. Akten des Symposiums. Heidelberg, 19–21. April 2012. Deutscher Terminologie-Tag e.V. Köln: SDK Systemdruck Köln GmbH & Co. KG. 101–118.

- Sermann E., Tamás D. 2010. Hogyan definiálhatjuk a fordítói adatbázist? Egy olasz és egy spanyol fordítói terminológiai adatbázis vizsgálata. In: Károly K., Fóris Á. (szerk.) *Nyelvek találkozása a fordításban. Doktori kutatások Klaudy Kinga tiszteletére*. Budapest: ELTE Eötvös Kiadó. 101–115.
- Tamás D. M., Sermann E. 2019a. Evaluation System for Online Terminological Databases. *Terminologija* 26. 24–46. (http://ki.lt/wp-content/uploads/2020/03/Terminologija_26_maketas.pdf).
- Tamás D.M., Sermann E. 2019b. Elemzési szempontrendszer terminológiai adatbázisokhoz. *Fordítástudomány* 21. évfolyam 2. szám, 46–62. (<https://doi.org/10.35924/fordtud.21.2.4>).
- Tamás D. M. 2020. Együttműködés a Szellemi Tulajdonjog Világszervezetével: a Budapesti Műszaki és Gazdaságtudományi Egyetemen végzett szabadalmi terminológiai egyetemi hallgatói projekt tapasztalatairól. *Magyar Jogi Nyelv* IV. évf. 2. szám. 26–41. (<https://joginyelv.hu/egyuttmukodes-a-szellemi-tulajdonjog-vilagszervezetével-a-budapesti-muszaki-es-gazdasagtudomanyi-egyetemen-vegzett-szabadalmi-terminologiai-egyetemi-hallgatoi-projekt-tapasztalatairól/>)
- Ugrin Zs. 2019. A terminuskapcsolatok és ontológiai rendszerek fedezése a fordítóképzésben. *Fordítástudomány* XXI. évf. 1. sz. 89–103.
- Valentini, C., Westgate, G., Rouquet, P. 2016. The PCT Termbase of the World Intellectual Property Organization. Designing a database for multilingual patent terminology. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* Vol. 22. No. 2. 171–200. (<https://doi.org/10.1075/term.22.2.02val>).
- Wright, S. E., Budin, G. 1997. *Handbook of Terminology Management*, Vol. 1. Amsterdam, Philadelphia: John Benjamins Publishing Company.
- Wüster, E. (Bauer, L. Hrsg.) 1985. *Einführung in die Allgemeine Terminologielehre und terminologische Lexikographie*. (2. Aufl.) Copenhagen: HEPU-tryk.

Források

A Szellemi Tulajdon Világszervezetét létrehozó 1967. július 14-i Stockholmi Egyezmény. Szabadalmi Együttműködési Szerződés (1970). (Patent Cooperation Treaty). Nemzetközi Szabadalmi Osztályozásra vonatkozó Strasbourgi Megállapodás (1971).

Szabványok

ISO 704:2009 - *Terminology work – Principles and methods*.

ISO 1087-1:2000 - *Terminology work – Vocabulary – Part 1: Theory and application*.

Elektronikus források

<https://www.sztnh.gov.hu/hu> (Letöltés ideje: 2020. 08. 20.)

<https://www.sztnh.gov.hu/hu/patent-cooperation-treaty-szabadalmi-egyuttmukodesi-szerzodes> (Letöltés ideje: 2020. 08. 22.)

<https://www.sztnh.gov.hu/hu/nemzetkozi-szabadalmi-osztalyozas-nszo-international-patent-classification-ipc> (Letöltés ideje: 2020. 08. 21.)

<https://www.wipo.int/about-wipo/en/> (Letöltés ideje: 2020. 08. 25.)

<https://www.wipo.int/classifications/ipc/en/> (Letöltés ideje: 2020. 08. 20.)

<https://wipopearl.wipo.int/en/conceptmap> (Letöltés ideje: 2020. 08. 20.)

<http://www.wipo.int/ipstats/en/> (Letöltés ideje: 2020. 08. 21.)

<https://www.wipo.int/PATENTSCOPE/en/> (Letöltés ideje: 2020. 08. 22.)

<https://www.wipo.int/pct/en/texts/glossary.html> (Letöltés ideje: 2020. 08. 21.)

<https://www.wipo.int/reference/en/wipopearl/> (Letöltés ideje: 2020. 08. 23.)

<https://www.wipo.int/reference/en/wipopearl/guide.html> (Letöltés ideje: 2020. 08. 20.)