

A kutatási terv közérdekű adatainak kivonata¹ beavatkozással nem járó vizsgálatok² számára³

A kitöltött nyomtatvány adatait az etikai véleményt adó Regionális Kutatásetikai Bizottságnak korlátozás nélkül hozzáférhetővé kell tennie bárki számára.

A kutatás-fejlesztési tevékenység során létrejövő szellemi javakat Magyarországon több törvény is védi.⁴ Ugyanakkor a Helsinkai Nyilatkozat 16. pontja, az Ovideoi Egyezményt hatályba léptető 2002. évi VI. törvény, és az orvosi kutatások végzéséről szóló miniszteri rendelet az emberen végzett orvosi kutatások etikus folytatása érdekében megkövetelik az etikai bizottságoktól, hogy a közvéleményt tájékoztassák az általuk véleményezett kutatások fontosabb adatairól. A közvélemény tájékoztatásának célja: az etikai bizottság munkájának nyilvánossága, a kutatások alanyai alapvető emberi jogainak biztosítása.

A 2007. III. 10-től hatályos 1/2007. (I. 24.) EüM rendelettel módosított 23/2002. (V. 9.) EüM rendelet szerint az alább felsorolt, a kutatási tervben megtalálható adatok közérdekű adatok, amelyeket bárki korlátozás nélkül megismerhet. Kérjük, hogy a szellemi alkotások oltalmának védelmét is szem előtt tartva, a nem nyilvános kutatási terv alapján töltsék ki ezt a táblázatot. A közvélemény és az alanyok tisztességes, lényegre törő tájékoztatását tartsa elsődleges szempontnak. A kutatási terv szakmai-etikai jóváhagyása után, az etikai bizottság a saját honlapján minden érdeklődő számára közzé teszi az itt megadott közérdekű adatokat. **Szakmai vagy szolgálati titoknak minősülő, illetve a kutatás érdekeit veszélyeztető adatot ne közöljön!**

A téma megnevezése (nem kell, hogy megegyezzen a kutatási protokoll címével)

Információkinyerés és szövegelemzés szöveges radiológiai leletek alapján

A kérelem iktatási száma: 34/2020-SZTE

A kérelmező neve, munkaköre és beosztása: Dr. Vidács László, tudományos főmunkatárs, Szegedi Tudományegyetem, MTA-SZTE Mesterséges Intelligencia Kutatócsoport és Szoftverfejlesztés Tanszék

¹ A 23/2002. (V. 9.) számú EüM rendelet 20/S. § (1) bekezdése szerint az itt felsorolt, a kutatási tervben is megtalálható adatok közérdekű nyilvános adatok, amelyeket az etikai bizottság a honlapján köteles nyilvánosságra hozni. Kérjük, hogy a szellemi alkotások oltalmának védelmét szem előtt tartva, a nem nyilvános kutatási terv alapján készítsék el ezt az adatlapot.

² A 23/2002. (V. 9.) számú EüM rendelet 20/B. § g) és h) pontjai szerint:

g.) *beavatkozással járó vizsgálat (interventional trial)*: fizikai beavatkozással járó orvostudományi kutatás és minden olyan beavatkozással járó kutatás, amely a vizsgálati alany lelki egészségére nézve kockázattal jár

h.) *beavatkozással nem járó vizsgálat (non-interventional trial)*: emberen végzett, a g) pont alá nem tartozó orvostudományi kutatás: 1/2007. (I.24) Eü.M. rendelet

³ Ez a nyomtatvány a 23/2002. (V. 9.) számú EüM rendelet 20/S. § (1) bekezdéseinek 2008. szeptember 1-jén hatályos szövege alapján készült.

⁴ A találmányok szabadalmi oltalmáról szóló 1995. évi XXXIII. törvény, a szerzői jogról szóló 1999. évi LXXVI. törvény.

1. A kutatás célja

A kutatás során anonimizált, szabad megfogalmazású radiológiai leletek szövegéből kívánunk természetesnyelv-feldolgozást alkalmazva (NLP), automatikus információkinyerést és szövegértelmezést végezni. Az elképzelés működőképességét első megközelítésben gerinc MR leletek feldolgozásával kívánjuk bemutatni.

2. A kutatás megszervezése és módszerei

A kutatásban a Szegedi Tudományegyetem Szoftverfejlesztés tanszék és az MTA-SZTE Mesterséges Intelligencia kutatócsoport informatikus és nyelvész szakemberei, illetve radiológus orvosok vesznek részt. A kutatáshoz a Klinikai Központ részéről anonimizált leletekre van szükség, melyeket egy előfeldolgozást követően radiológusok annotálnak egy előre meghatározott útmutató alapján. Az annotált leletek képezik a további szövegbányászati módszerek alapját. A leletek szövegének automatikus értelmezését gépi tanulási, így többek között mélytanulási (deep learning) eljárással végezzük.

3. A tervezett kutatás szakirodalmi alapjai

Ugyan próbálkozások folyamatosan történnek az elektronikus leletezés strukturáltabbá, szabványosabbá tételére, a tapasztalat azt mutatja, hogy mind a radiológusok, mind a szakorvosok előnyben részesítik a szabad megfogalmazású leletek készítését. Ez vitathatatlanul elősegíti a leletek nyelvi gazdagságát és szabatos megfogalmazását, azonban jelentősen megnehezíti a leletekből történő információkinyerést, illetve a leletezés folyamatának minőségbiztosítását. Ez a kihívás kiváló lehetőséget ad klinikai környezetre specializált, mesterséges intelligenciával támogatott, természetesnyelv-feldolgozó rendszerek fejlesztésére. Az elmúlt évtizedben a diagnózissegéd [1], [2], [3] és minőségbiztosítást végző [4], [5], [6], [7] alkalmazások fejlesztése folyamatosan bővült. Az utóbbi évek mesterséges intelligenciával támogatott rendszereinek fejlesztését, valamint a jelentősebb előrelépéseket több összefoglaló cikk is tárgyalja [8], [9], [10], [11], [12], [13]. A terület látványos fejlődése ellenére a gerincről készült radiológiai leletek mesterséges intelligenciával történő feldolgozása a nemzetközi szakirodalomban is ritkának mondható. Példaként említhető Tan és munkatársainak munkája, mely során 26 alsóháti fájdalomra utaló orvosi megállapítás, radiológiai leletekben történő azonosítását végezték egy gépi tanuláson alapuló és egy szabályalapú rendszer összehasonlítása során [14]. A szerzők egy másik munkájukban szabályalapú rendszert fejlesztettek 1-es típusú Modic véglemezelváltozások felismerésére [15]. Wang és szerzőtársai csonttritkulásból eredő törések, többek között gerinctörések radiológiai leletekben történő felismerésére fejlesztettek szabályalapú NLP alkalmazást [16]. Érdemes megemlíteni, hogy a magyar nyelvű radiológiai leleteken tudomásunk szerint még nem végeztek az általunk javasolt kutatáshoz hasonlót. Megjegyzendő azonban, hogy magyar nyelvű orvosi leleteken is történtek már kutatások, így például fejlesztettek rendszert orvosi szövegek automatikus szegmentálására, valamint az orvosi rövidítések automatikus kezelésére [17, 18], elgépelések javítására [19], illetve orvosi szövegekben előforduló releváns kifejezések azonosítására [20].

Kutatásunk a magyar nyelvű klinikai szövegekből történő információkinyerés egy következő mérföldköve lehet, mely egyelőre szöveges gerinc MR leletek elemzésére szorítkozna.

Irodalomjegyzék

- [1] Pham, A.D., Névéol, A., Lavergne, T., Yasunaga, D., Clément, O., Meyer, G., Morello, R., Burgun, A.: Natural Language Processing of Radiology Reports for the Detection of Thromboembolic Diseases and Clinically Relevant Incidental Findings. *BMC Bioinformatics* **15**(1) (2014) 266
- [2] Rink, B., Roberts, K., Harabagiu, S., Scheuermann, R.H., Toomay, S., Browning, T., Bosler, T., Peshock, R.: Extracting Actionable Findings of Appendicitis from Radiology Reports Using Natural Language Processing. *AMIA Joint Summits on Translational Science Proceedings. AMIA Joint Summits on Translational Science* **2013** (2013) 221
- [3] Solti, I., Cooke, C.R., Xia, F., Wurfel, M.M.: Automated Classification of Radiology Reports for Acute Lung Injury: Comparison of Keyword and Machine Learning Based Natural Language Processing Approaches. In: *Proceedings - 2009 IEEE International Conference on Bioinformatics and Biomedicine Workshops, BIBMW 2009. Volume 2009., NIH Public Access* (2009) 314–319
- [4] Raja, A.S., Ip, I.K., Prevedello, L.M., Sodickson, A.D., Farkas, C., Zane, R.D., Hanson, R., Goldhaber, S.Z., Gill, R.R., Khorasani, R.: Effect of Computerized Clinical Decision Support on the Use and Yield of CT Pulmonary Angiography in the Emergency Department. *Radiology* **262**(2) (2012) 468–474
- [5] Ip, I.K., Morteale, K.J., Prevedello, L.M., Khorasani, R.: Focal Cystic Pancreatic Lesions: Assessing Variation in Radiologists' Management Recommendations. *Radiology* **259**(1) (2011) 136–41
- [6] Siström, C.L., Dreyer, K.J., Dang, P.P., Weilburg, J.B., Boland, G.W., Rosenthal, D.I., Thrall, J.H.: Recommendations for Additional Imaging in Radiology Reports: Multifactorial Analysis of 5.9 Million Examinations. *Radiology* **253**(2) (2009) 453–61
- [7] Dang, P.A., Kalra, M.K., Blake, M.A., Schultz, T.J., Stout, M., Lemay, P.R., Freshman, D.J., Halpern, E.F., Dreyer, K.J.: Natural Language Processing Using Online Analytic Processing for Assessing Recommendations in Radiology Reports. *Journal of the American College of Radiology* **5**(3) (2008) 197–204
- [8] Wang, Y., Wang, L., Rastegar-Mojarad, M., Moon, S., Shen, F., Afzal, N., Liu, S., Zeng, Y., Mehrabi, S., Sohn, S., Liu, H.: Clinical Information Extraction Applications: A Literature Review (2018)
- [9] Pons, E., Braun, L.M., Hunink, M.G., Kors, J.A.: Natural Language Processing in Radiology: A Systematic Review. *Radiology* **279**(2) (2016) 329–343
- [10] Ford, E., Carroll, J.A., Smith, H.E., Scott, D., Cassell, J.A.: Extracting Information from the Text of Electronic Medical Records to Improve Case Detection: A Systematic Review. *Journal of the American Medical Informatics Association* **23**(5) (2016) 1007–1015
- [11] Cai, T., Giannopoulos, A.A., Yu, S., Kelil, T., Ripley, B., Kumamaru, K.K., Rybicki, F.J., Mitsouras, D.: Natural Language Processing Technologies in Radiology Research and Clinical Applications. *RadioGraphics* **36**(1) (2016) 176–191
- [12] Yim, W.w., Yetisgen, M., Harris, W.P., Kwan, S.W.: Natural Language Processing in Oncology. *JAMA Oncology* **2**(6) (2016) 797
- [13] Meystre, S.M., Savova, G.K., Kipper-Schuler, K.C., Hurdle, J.F.: Extracting Information from Textual Documents in the Electronic Health Record: A Review of Recent Research. *Yearbook of Medical Informatics* (2008) 128–44
- [14] Tan, W.K., Hassanpour, S., Heagerty, P.J., Rundell, S.D., Suri, P., Huhdanpaa, H.T., James, K., Carrell, D.S., Langlotz, C.P., Organ, N.L., Meier, E.N., Sherman, K.J., Kallmes, D.F., Luetmer, P.H., Griffith, B., Nerenz, D.R., Jarvik, J.G.: Comparison of Natural Language Processing Rules-Based and Machine-Learning Systems to Identify Lumbar Spine Imaging Findings Related to Low Back Pain (2018)
- [15] Huhdanpaa, H.T., Tan, W.K., Rundell, S.D., Suri, P., Chokshi, F.H., Comstock, B.A., Heagerty, P.J., James, K.T., Avins, A.L., Nedeljkovic, S.S., Nerenz, D.R., Kallmes, D.F., Luetmer, P.H., Sherman, K.J., Organ, N.L., Griffith, B., Langlotz, C.P., Carrell, D., Hassanpour, S., Jarvik, J.G.: Using Natural Language Processing of Free-Text Radiology Reports to Identify Type 1 Modic Endplate Changes. *Journal of Digital Imaging* **31**(1) (2018) 84–90
- [16] Wang, Y., Mehrabi, S., Sohn, S., Atkinson, E., Amin, S., Liu, H.: Automatic Extraction of Major Osteoporotic Fractures from Radiology Reports using Natural Language Processing. In: *Proceedings -*

2018 IEEE International Conference on Healthcare Informatics Workshops, ICHI-W 2018, IEEE (2018) 64–65

[17] Orosz, Gy., Novák, A., Prószéky, G.: Hybrid Text Segmentation for Hungarian Clinical Records. In Castro, F., Gelbukh, A., González, M., eds.: *Advances in Artificial Intelligence and Its Applications: 12th Mexican International Conference on Artificial Intelligence, MICAI 2013*, Mexico City, Mexico, November 24–30, 2013, Proceedings, Part I, Berlin, Heidelberg, Springer Berlin Heidelberg, Springer Berlin Heidelberg (2013) 306–317

[18] Siklósi, B., Novák, A.: Rec. et exp. aut. Abbr. mnyelv. KLIN. szövegben – Rövidítések Automatikus Felismerése és Feloldása Magyar Nyelvű Klinikai Szövegekben. In Tanács, A., Varga, V., Vincze, V., eds.: *X. Magyar Számítógépes Nyelvészeti Konferencia*, Szeged, Szegedi Tudományegyetem, Informatikai Tanszékcsoport (2014) 167–176

[19] Siklósi, B., Orosz, Gy., Novák, A., Prószéky, G.: Automatic Structuring and Correction Suggestion System for Hungarian Clinical Records. In De Pauw, G., de Schryver, G.M., Forcada, M.L., M. Tyers, F., Waiganjo Wagacha, P., eds.: *8th SaLTMiL Workshop on Creation and Use of Basic Lexical Resources for LessResourced Languages*, Istanbul (2012) 29–34

[20] Siklósi, B., Novák, A.: Nem Felügyelt Módszerek Alkalmazása Releváns Kifejezések Azonosítására és Csoportosítására Klinikai Dokumentumokban. In Tanács, A., Varga, V., Vincze, V., eds.: *XI. Magyar Számítógépes Nyelvészeti Konferencia (MSZNY 2015)*, Szeged, Szegedi Tudományegyetem Informatikai Tanszékcsoport, Szegedi Tudományegyetem Informatikai Tanszékcsoport (2015) 237–248

4. A kutatásba bevonni kívánt vizsgálati alanyok száma, illetve köre, neme, életkora

A kutatás korábbi radiológiai leletek szövegének vizsgálatán alapul anoním módon, a feldolgozni tervezett leletek száma ezres nagyságrendű.

5. Retrospektív, beavatkozással nem járó vizsgálatok esetében, amikor a 23/2002. (V. 9.) számú EüM rendelet 20/Q. §-ának alkalmazására kerül sor, a nyilvános adatvédelmi tájékoztatás⁵ (A tájékoztatás – különösen statisztikai vagy tudományos célú adatkezelés esetén – megtörténhet az adatgyűjtés tényének, az érintettek körének, az adatgyűjtés céljának, az adatkezelés időtartamának és az adatok megismerhetőségének mindenki számára hozzáférhető módon történő nyilvánosságra hozatalával, ha az egyénre szóló tájékoztatás lehetetlen vagy aránytalan költséggel járna.)

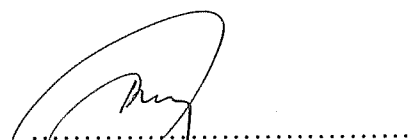
A vizsgálatról tájékoztatást a Radiológiai Klinika és az MTA-SZTE Mesterséges Intelligencia Kutatócsoport honlapján tesszük közzé.

Nyilatkozom, hogy a fenti adatok nem sértik a kutatásnak a szellemi alkotások védelmére vonatkozó érdekeit és nem tartalmaznak szakmai- vagy szolgálati titkot, illetve a kutatás érdekeit veszélyeztető adatot. A fenti adatokat bárki, korlátozás nélkül megismerheti. Tudomásul veszem, hogy jóváhagyás után az RKEB a közérdekű adatokat a honlapján közzé teszi.

Szeged, 2019. december 17.


.....
DR. VIDÁCS LÁSZLÓ


.....
DR. GYIMÓTHY TIBOR


.....
DR. KINCSES ZSIGMOND
TAMÁS

⁵ A beavatkozással nem járó, retrospektív, statisztikai vizsgálatok esetén – ahol az egyénre szóló tájékoztatás lehetetlen vagy aránytalanul nagy költséggel járna a 23/2002. (V. 9.) számú EüM. rendelet 20/Q. §-a alapján el lehet tekinteni a vizsgálati alany, illetve kiskorú, cselekvőképtelen vagy korlátozottan cselekvőképes személy esetén a törvényes képviselő tájékoztatásától és a beleegyező nyilatkozat beszerzésétől.

Az alább idézett adatvédelmi törvény 6. § (4) bekezdése szerint, ilyen esetben a tájékoztatás az adatgyűjtés tényének, az érintettek körének, az adatgyűjtés céljának, az adatkezelés időtartamának és az adatok megismerhetőségének mindenki számára hozzáférhető módon történő nyilvánosságra hozásával történik. Ezeket az adatokat kérjük közölni.

A Személyes adatok védelméről és a közérdekű adatok nyilvánosságáról szóló 1992. évi LXIII. törvény 6. § (1) *Az érintettel az adat felvétele előtt közölni kell, hogy az adatszolgáltatás önkéntes vagy kötelező. Kötelező adatszolgáltatás esetén meg kell jelölni az adatkezelést elrendelő jogszabályt is.*

(2) *Az érintettet – egyértelműen és részletesen – tájékoztatni kell az adatai kezelésével kapcsolatos minden tényről, így különösen az adatkezelés céljáról és jogalapjáról, az adatkezelésre és az adatfeldolgozásra jogosult személyéről, az adatkezelés időtartamáról, illetve arról, hogy kik ismerhetik meg az adatokat. A tájékoztatásnak ki kell terjednie az érintett adatkezeléssel kapcsolatos jogaira és jogorvoslati lehetőségeire is.*

(4) *A tájékoztatás – különösen statisztikai vagy tudományos (ideértve a történelmi kutatásokat is) célú adatkezelés esetén – megtörténhet az adatgyűjtés tényének, az érintettek körének, az adatgyűjtés céljának, az adatkezelés időtartamának és az adatok megismerhetőségének mindenki számára hozzáférhető módon történő nyilvánosságra hozatalával, ha az egyénre szóló tájékoztatás lehetetlen vagy aránytalan költséggel járna.*