
SMURTO KONTROLĖS LIETUVIŠKAME SAITYNE SĄRYŠIS SU SEMANTINIŲ IR KALBOS TECHNOLOGIJŲ PLĖTRA

Darius Amilevičius

*Mykolo Romerio universiteto Viešojo saugumo fakulteto Humanitarinių mokslų katedra
Putvinskio g. 70, LT-44211 Kaunas
Telefonas 30310
El.paštas: d.amilevicius@mruni.eu*

Anotacija. Elektroninėje erdvėje sparčiai plinta nekontroliuojamas elektroninis smurtas. Elektroninės kontrolės objekto – informacijos – kiekis saityne irgi sparčiai didėja. Esami skaičiavimo metodai jau nesugeba jo apdoroti. Nors vis didėja informacijos dalis, perduodama vaizdu ir garsu, bet tekstas išlieka pagrindine informacijos perdavimo forma. Didelius kiekius tekstinės informacijos ar kitokių nestruktūrizuotų duomenų analizuojančios teisėsaugos bei viešojo saugumo užtikrinimo institucijos susiduria su problema, vykdydamos informacijos lietuvių kalba paiešką ir analizę – nėra veiksmingų prasminės (semantinės) teksto paieškos ir analizės priemonių, pritaikytų dirbti su lietuvių kalbos tekstais ir duomenimis. Reikalingos kokybiškai naujos, semantinės ir kalbų technologijos, kurios mokėtų analizuoti ir apibendrinti informaciją pagal jos prasmę. Atlikus tyrimą nustatyta, kad Lietuvos mokslininkai jau pajėgūs sukurti ir įdiegti semantinių ir kalbos technologijų naujoves, būtinas teisėsaugos bei viešąjį saugumą užtikrinančioms institucijoms. Tačiau tai padaryti kliūdo vėluojančios valstybės strateginės programos ir priemonės, bendros kalbos technologijų išteklių ir įrankių vystymo koncepcijos nebuvimas, nepakankamos valstybės investicijos į semantinio pasaulinio tinklo technologijas, ribotas bendradarbiavimas tarp privataus ir viešojo sektoriaus, bei akademinės bendruomenės.

Pagrindinės sąvokos: elektroninio smurto kontrolė, semantinės technologijos, kalbos technologijos, saugumas elektroninėje erdvėje.

ĮVADAS

Pasaulyje, Europos Sąjungoje ir Lietuvoje vyksta svarbūs transformacijos procesai. Nuo 2000 m. industrinę erą pakeitė informacinė, kuri pasižymi ne tik sparčiai gausėjančiu informacijos kiekiu ir daugelio viešųjų bei verslo paslaugų ir procesų perkėlimu į elektroninę erdvę, bet ir naujų ekonominių sistemų bei socialinių santykių formavimusi. Technologinė pažanga sudarė sąlygas ir neigiamiems procesams. Tiek realiame gyvenime, tiek elektroninėje erdvėje įsigalėjusio smurto mastai pralenkė ne tik galimybes išaiškinti ir nubausti nusikaltėlius, bet ir registruot. Be to, smarkiai pakito pati nusikaltimų sąvoka: visuomenė smurto kartelę pakelė labai aukštai. Smulkiais, dėmesio nevertais nusikaltimais tapo kišenių apšvarinimas, nosies suknežinimas patamsyje, grasinimai, smurtas artimųjų tarpe, apvogti automobiliai, išmalti langai, nepageidaujami elektroniniai laiškai, elektroninės patyčios, elektroninis priekabiavimas. Dažnai jie raminančiai pavadinti buitinais nusikaltimais.

Temos aktualumas. Informacijos kiekis sparčiai didėja, esami skaičiavimo metodai jau nesugeba jo apdoroti. Nors vis didėja informacijos dalis, perduodama vaizdu ir garsu, bet tekstas išlieka pagrindine informacijos perdavimo forma. Todėl reikalingos kokybiškai naujos, semantinės ir kalbų technologijos, kurios mokėtų analizuoti ir apibendrinti informaciją pagal jos prasmę. Tai leistų piliečiams naudotis virtualios erdvės paslaugomis jų gimtąja kalba, padėtų išsaugoti gimtąją kalbą ir tautinį tapatumą globalizacijos ir skaitlingos migracijos kontekste. Be to, tai suteiktų teisėsaugos institucijoms veiksmingus technologinius sprendimus ir priemones elektroninio smurto kontrolei, stiprinant visuomenės saugumą ir viešąją tvarką elektroninėje erdvėje.

Tyrimo tikslas – atskleisti semantinių ir kalbos technologijų taikymo galimybes, vykdant elektroninio smurto lietuviškame saityne veiksmingą kontrolę.

Uždaviniai:

- Ištirti elektroninio smurto lietuviškame saityne esamą situaciją, nustatyti ateities prognozes.
- Ištirti semantinių ir kalbos technologijų plėtrą Lietuvoje ir užsienyje.

Išryškinti semantinių ir kalbos technologijų taikymo galimybes smurto kontrolei lietuviškame saityne.

Tyrimo objektu pasirinktas elektroninio smurto kontrolė lietuviškame saityne ir jos optimizavimo galimybės, pasitelkiant semantinių bei kalbinių technologijų siūlomus sprendimus.

Tyrimo metodai. Pirminių duomenų rinkimo metodas - apklausa žodžiu. Buvo apklausti dešimt specialiųjų tarnybų vadovų ir specialistų, siekiant atskleisti semantinių ir kalbų technologijų taikymo galimybes, vykdant elektroninio smurto kontrolę. Dokumentų analizės metodas taikytas siekiant įvertinti elektroninio smurto lygį lietuviškame saityne ir palyginti tarpusavyje įvairius semantinių ir kalbos technologijų sprendimus bei taikymus. Taip pat naudoti mokslinės literatūros analizė, interpretacijos, sintezės, vertinimo metodai.

SMURTAS SAITYNE: ESAMA SITUACIJA IR PERSPEKTYVOS PROGNOZĖ

Andrew P. Morriss nuomone, elektroninė erdvė prilygintina „laukiniams Vakarams“, kur žmogui atsiveria neribotos galimybės, bet kur nevaržomai plinta smurtas ir prievarta, o niekadėjai nesulaukia pelnytų bausmės¹. Anne S.Y. Cheung teigia, kad nors internetas

¹ Morriss, A. P. The Wild West Meets Cyberspace. *The freeman: ideas on liberty* [interaktyvus]. 1998, 48(7) [žiūrėta 2011.09.30]. <<http://www.thefreemanonline.org/featured/the-wild-west-meets-cyberspace/print/>>.

daugeliui žmonių sudarė iki tol neturėtas prieinamumo, betarpiškumo, anonimiškumo galimybes, bet kartu tapo iššūkiu žmogaus teisėms, asmens ir visuomenės saugumui, nes prasidėjo precedento neturinčios smurto apraiškos (elektroninis priekabiavimas, elektroninis šmeižtas ir kt.)². Šiais laikais elektroninėje erdvėje vis daugiau informacijos perduodamas garsu ir vaizdu, bet tekstas vis dar lieka pagrindine informacijos perdavimo forma. Internetu ir elektroninėje erdvėje egzistuoja ir proporcingai nuo vartojimo didėja įvairių grėsmių ir pavojų, atsitikimų, incidentų tikimybė, kuri gali paliesti beveik kiekvieną interneto vartotoją ir juo labiau paauglį³. Valstybės mastu šios grėsmės yra įvardintos, nors ir nepilnai, joms numatytos prevencinės ir baudžiamosios priemonės, bet kontrolė nevykdoma.

Galimybė platinti ar net skatinti smurtą žiniasklaidos priemonėmis ir naujomis technologijomis bei būtinybė žmones nuo to apsaugoti yra problema, apie kurią diskutuojama nuo tada, kai atsirado ryšių technologijos, ir ši diskusija toli gražu nėra baigta, o greičiau tik atsinaujina kartu su kiekvienu technologiniu išradimu. Daugelis pavojų, kuriuos kelia IRT technologijos, gresia visiems žmonėms, tačiau vaikai ir paaugliai yra ypač pažeidžiami ne tik dėl savo jaunumo, bet ir dėl to, kad naujos technologijos jų gyvenime yra kur kas įtakingesnės. Ko gero, svarbiausia iš problemų, siejančių naujas technologijas su smurtu, yra baimė, kad lengviau pažeidžiami žmonės gali tapti vadinamųjų „internetinių grobuonių“ (angl. *online predators*), ieškančių aukų, kurias jie gali izoliuoti ir išnaudoti, taikiniai. Žinomas interneto anonimiškumas sukūrė galimybes žmonėms „užsimaskuoti“ ir vaidinti internetinę asmenybę, taip lengviau bendraujant su vaikais, paaugliais ir pažeidžiamais suaugusiais pokalbių kambariuose ir tinklaraščių svetainėse. Šiam procesui apibūdinti vartojami tokie terminai, kaip „persekiojimas“ (angl. *stalking*) ir „apdorėjimas“ (angl. *grooming*), atspindintys įtikinantį ir dominuojantį šių kontaktų pobūdį⁴.

Dėl straipsnio apimties neįmanoma pateikti visų formų smurto elektroninėje erdvėje analizės. Apsiribosime šiuo metu Lietuvoje pakankamai ištirtos ir ypač skaudžios elektroninio smurto formos - elektroninių patyčių – statistinių duomenų analize. Objektyvumo dėlei reikia pripažinti, kad, neskaitant elektroninių saugumo pažeidimų statistikos, pateikiamos Lietuvos Respublikos nacionalinis elektroninių ryšių tinklų ir informacijos saugumo incidentų tyrimo

² Cheung, A. S. Y. A Study of Cyber-Violence and Internet Service Providers' liability: Lessons from China. *Pacific Rim Law & Policy Journal*. 2009, 18(2): 345.

³ Ruškus J.; Žvirdauskas; Kačėnuskaitė, V.; Žvirdauskienė R. Paauglių elgsena elektroninėje erdvėje [interaktyvus]. Vilnius: Švietimo informacinių technologijų centras [žiūrėta 2011.09.30]. 2009: 13 psl <http://www.draugiskasinternetas.lt/repository/dokumentai/11_03_24/tyrimas%20elgsena%20elektronineje%20aplinkoje.pdf>.

⁴ Kane, J.; Portin, P. Smurtas ir technologijos [interaktyvus]. Briuselis [žiūrėta 2011.09.30]. 2008: 7-9. <http://ec.europa.eu/justice_home/daphnetoolkit/files/others/booklets/07_daphne_booklet_7_lt.pdf>.

padalinys (CERT-LT) tinklalapyje⁵, kitos elektroninio smurto apraiškų lietuviškame saityne tyrimai fragmentuoti ir nepakankamai informatyvūs.

Elektroninės patyčios yra viena iš psichologinio smurto formų, agresyvios elgsenos subkategorija. Tai vis labiau plintantis fenomenas, tapęs tarptautine problema. Dažniausiai elektroninės patyčios nėra atsitiktinis informacijos perdavimas. Tai kartojasi dažnai ir nukentėjusiajam yra sunku apsiginti⁶. Jį skatina didėjantis internetinių socialinių tinklapių ir portalų tarp vaikų ir paauglių populiarėjimas. Elektroninės patyčios – tai elektroninių komunikacinių technologijų naudojimas, siekiant tyčia užgauti, išgąsdinti, pažeminti individą ar grupę, paskleisti melagingą, įžeidžiančią informaciją ir kt. Anksčiau vaikai erzindavo vieni kitus žaisdami žaidimų aikštelėse, dabar jie tai daro elektroninėse svetainėse. Skriaudėjai pateikia aukų asmeninę informaciją ir ji tampa prieinama visiems. Informacijos plitimas elektroninėje erdvėje sukelia didelę grėsmę aukų privatumui. Informacija gali būti persiųsta tūkstančiams adresatų per keletą sekundžių ir to negalima sukontroliuoti. Patiriantis elektronines patyčias atsiduria nemalonioje situacijoje ir nežino, kaip apsiginti. Jis gali: gauti užgaulius, grasinančius elektroninius laiškus ar SMS žinutes, mobiliuoju telefonu gauti trikdančią nuotrauką, sulaukti grasinimų pokalbių svetainėje, pokalbių svetainėje perskaityti apie save melagingus gandus, būti išmestas iš pokalbių svetainės, būti pažemintas pamatęs išsiuntinėtą suklastotą savo nuotrauką, būti įskaudintas, pamatęs paskleistą asmeninio pobūdžio informaciją apie save, nemaloniai pasijusti sužinojęs, kad be leidimo pasinaudota asmeniniame kompiuteryje esančia informacija. Lietuvos Respublikos Socialinės apsaugos ir darbo ministerijos užsakymu atliktas tyrimas apie interneto, kompiuterinių žaidimų įtaką smurtui prieš vaikus parodė, kad tarp dažniausiai lankomų interneto puslapių vaikai nurodo įvairiausių žaidimų, pokalbių bei pažinčių svetainių internetinius adresus. Su amžiumi mažėja susidomėjimas kompiuteriniais žaidimais, kyla pažinčių svetainių, socialinių tinklų populiarumas⁷. Tokių tinklapių patrauklumą didina ne tik galimybės susirasti interesus atitinkančią medžiagą, bet ir jų interaktyvumas. Iš interneto imama, į internetą talpinama, čia pat komentuojama. Apklausoje aukščiausią reitingą gavo vaizdo ir garso mainų svetainės, trečioje vietoje – socialiniai tinklai. Dažniausiai vieną prie interneto mokinį laiko žaidimai bei

⁵ Lietuvos Respublikos nacionalinis elektroninių ryšių tinklų ir informacijos saugumo incidentų tyrimo padalinys (CERT-LT): metinė ir ketvirtinė incidentų statistika [interaktyvus]. Vilnius, 2011 [žiūrėta 2011.09.30]. <<https://www.cert.lt/statistika.html>>.

⁶ Smith P.K.; Morita, Y. *The Nature of School Bullying. A cross-national perspective*. In P.K. Smith et al (eds). London and New York: Routledge, 1999, p. 1.

⁷ Interneto, kompiuterinių žaidimų įtaka smurtui prieš vaikus [interaktyvus]. Vilnius: Socialinės apsaugos ir darbo ministerija, UAB „Factus Dominus“ [žiūrėta 2011.09.30]. 2005. <<http://www.draugiskasinternetas.lt/lt/adult/news?id=2003>>.

socialiniai tinklai. Žaidimai bei socialiniai tinklai yra dominuojančios interneto temos pagal ten praleistą laiką (vertinant paauglių interneto vartojimui skiriamas laiko sąnaudas pagal lytį, daugiausiai laiko socialiniams tinklams naršyti skiria mergaitės. Nepažįstamųjų laišakai yra potencialiai pavojingi paaugliams, kadangi paaugliai nepažįstamųjų laiškus vertina labiau teigiamai, kaip galimybes, nei neigiamai, kaip grėsmes. Tai liudija apie dar nepakankamai paauglių įsisąmonintą grėsmę, juolab žinant, kad paauglystėje vaikai tikisi naujų santykių, pažinčių, tuo pačiu, pripažinimo, supratimo, dalyvavimo. Paaugliai suvokia, kad, greta daugelio naudų, gaunamų virtualiu būdu jungiantis prie socialinių tinklų, informacinių ar pramogų šaltinių, šie šaltiniai gali tapti ir tam tikrų grėsmių ir pavojų šaltiniu. Tyrimą atlikę mokslininkai mano, kad jokios draudžiamos ar moralizavimo priemonės nepadės geriau suvokti interneto grėsmių – būtina diskusija⁸. Bet ar galime laukti ir tikėtis, kad paaugliai patys save apsaugos?

Reikšmingi elektroninių patyčių paplitimo Klaipėdos apskrities vyresniųjų moksleivių tarpe tyrimų rezultatai. Nustatyta, kad 30,3 % moksleivių per pastaruosius kelis mėnesius bent kartą patyrė elektronines patyčias internetu ar mobiliuoju telefonu. Patyčias vienodai dažnai patiria visų klasių ir abiejų lyčių moksleiviai. Daugiausiai patyčių 5-8 klasių moksleiviai patiria pokalbių svetainėse. Į klausimą „Ką tu darei, kai tave įskaudino?“. 34 procentai apklaustųjų pasirinko atsakymą „Niekam nesakiau“. Tik apie 5 procentus apklaustųjų apie patyčias informuoja mokyklos psichologą ar pedagogus. Tyčiojimas elektroninėje erdvėje yra susijęs su tradicinėmis patyčiomis. Tradicinių patyčių „aukos“ dažniau tampa ir elektroninių patyčių aukomis, tradiciškai besityčiojantys dažniau tampa „skriaudėjais“ ir elektroninėje erdvėje, taip pat tradicinių patyčių „aukos“ dažniau tampa „skriaudėjais“ elektroninėje erdvėje⁹. Lietuvoje atlikto tyrimo su 15–19 metų moksleiviais duomenys patvirtina užsienio mokslininkų iškeltą mintį, kad net ir vyresnėse klasėse agresijos problema yra aktuali ir ją sprendžiant turi būti atsižvelgta į tai, kad vyresnio mokyklinio amžiaus moksleivių keičiasi agresijos raiška, t.y. jaunuoliams bręstant agresija ne mažėja, bet keičiasi jos forma¹⁰.

Kalbant apie smurto elektroninėje erdvėje plitimą ateityje, būtina įvertinti ir informacijų technologijų sklaidą Lietuvoje. Tyrimai patvirtina, kad Lietuvos rinka instituciniame ir

⁸ Ruškus, J. *Paauglių elgsena elektroninėje, supra note 3*, p.

⁹ Elektroninių patyčių paplitimas Klaipėdos apskrities vyresniųjų moksleivių tarpe (tyrimo duomenys) [interaktyvus]. Klaipėdos universiteto Psichologijos katedra [žiūrėta 2011.09.30]. <http://www.ku.lt/pf/psichologija/doc/Tyrimo_duomenys_1.pdf>.

¹⁰ Pilkauskaitė Valickienė, R.; Raižienė, S.; Žukauskienė, R. Elektroninių Patyčių Paplitimas tarp Klaipėdos Apskrities Vyresniųjų Klasių Moksleivių. *Socialinis darbas*. 2009 m. Nr. 8(2): 119.

individualiame lygmenyse pasirengusi intensyviai informacinių technologijų sklaidai. Europos Komisijos „Europos skaitmeninės darbotvarkės“ ataskaitos duomenimis, Lietuva yra viena pirmaujančių Europos Sąjungos valstybių pagal šalies įmonių aktyvumą naudojantis informacinių technologijų privalumais atspindinčius rodiklius pvz.: įmonių, kurios internetu bendrauja su viešojo sektoriaus institucijomis (95 proc.)¹¹. Bendrą situaciją atspindi ir Informacinės Visuomenės Plėtros Komiteto prie Susisiekimo ministerijos užsakymu atliktų tyrimų rezultatai. 2011 m. pradžios duomenimis, internetu jau naudojasi 8 kartus daugiau Lietuvos gyventojų nei 2001 m. – 63,5 proc. šalies gyventojų ir net 100 proc. 15-19 metų amžiaus Lietuvos jaunimo. Beveik 60 proc. namų ūkių turi interneto ryšį, kad internete lietuviai dažniausiai naudojami paieškos sistemomis (tai daro 82 proc. interneto vartotojų), elektroniniu paštu (78 proc.), internetiniais portalais (71 proc.), interneto forumais ir pokalbių svetainėmis (59 proc.)¹². Europos Sąjungos skaitmeninės darbotvarkės tikslas iki 2015 m. pasiekti, kad reguliariai internetu besinaudojančių žmonių padaugėtų nuo 60 iki 75 proc.¹³. Lietuvos rinkotyros prognozes, kad 2020 m. Lietuvoje gyventojų, besinaudojančių internetu, skaičius išaugs iki 80 proc.

Apibendrinant, darytina išvada, kad smurtas elektroninėje erdvėje jau tapęs tarptautine problema. Intensyvi interneto ir informacinių technologijų sklaida gyventojų tarpe leidžia prognozuoti, kad nesant efektyvių poveikio priemonių, smurto atvejų elektroninėje erdvėje tik daugės. Auka dažnai nutyli apie patirtą elektroninį smurtą ir, dažniausia, negali pati apsiginti. Tai rodo, kad, siekiant užtikrinti piliečių saugumą internete, efektyvių prevencijos priemonių jau nepakanka. Būtinės veiksmingos elektroninės erdvės kontrolės priemonės.

SEMANTINIŲ IR KALBOS TECHNOLOGIJŲ PLĖTRA EUROPOS SAJUNGOJE IR LIETUVOJE

Semantinis saitynas – kartais vadinamas trečia saityno versija – vaidina svarbiausią vaidmenį, vystant saityno informacines technologijas ir paslaugas. Jis lemia naujų iššūkių formavimosi procesą, kai informacijos prasmė leidžia mašinai suprasti saityno turinį ir imituoti žmogaus intelektą, atliekant daug varginančių užduočių, tokių kaip informacijos paieška, pasidalinimas informacija, informacijos iš kelių šaltinių susiejimas. Iki šiol mašina

¹¹ Europe's Digital Competitiveness Report - ICT Country Profiles [interaktyvus]. [žiūrėta 2011.09.30].

<http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=668>

¹² Informacinės Visuomenės Plėtros Komiteto prie Susisiekimo ministerijos statistika [interaktyvi]. Vilnius: 2011 [žiūrėta 2011.09.30]. <<http://www.ivpk.lt/lt/ltvm/statistika>>.

¹³ Europe's Digital Competitiveness Report [interaktyvus]. [žiūrėta 2011.09.30].

<http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=667>.

negalėjo pilnai atlikti šias užduotis be žmogaus įsikišimo, nes pirmosios ir antrosios kartos saityno informacijos prasmė suprantama tik žmogui. Mašina iki šiol sugeba tik pernešti ją pagal nuorodą ir nustatyta forma pateikti adresatui. Saityno kūrėjas Timas Berners-Lee dar 2001 m. prognozavo, kad kompiuteriai ateityje bus pajėgūs pilnai analizuoti visą saityno informaciją – turinį, nuorodas, transakcijas tarp žmogaus ir mašinos. Jis teigė, kad tai padaryti leis tuo metu dar neegzistavęs semantinis saitynas, kurio dėka kasdienės, rutininės operacijas – prekybos, biurokratijos, eilinio vartotojo buities ir t.t. - atliks mašina, bendraujanti su kita mašina¹⁴.

Semantinis pasaulinis tinklas jau veikia. Didžiosios Europos Sąjungos šalys skubiai kuria semantines informacinių ir ryšių technologijų (toliau – IRT) infrastruktūras. Šiuo klausimu toliausiai pažengusi Suomija, kurios Vyriausybė jau 2001 m. pastebėjo, kad investicijos į semantines IRT infrastruktūras prilygintinos investicijoms į kelių infrastruktūrą. Verslas niekada neinvestuoja į kelių infrastruktūros vystymą, bet greitai sukuria būdus, kaip ją efektyviai panaudoti, didindamas šalies Bendrąjį vidaus produktą, kurdamas darbo vietas, sumokėdamas daugiau mokesčių. Todėl Suomijos semantinė IRT infrastruktūra išvystyta geriausiai Europos Sąjungoje: veikia semantiniai portalai, teikiamos semantinės IRT paslaugos. Infrastruktūros vystymą užsako valdžia, todėl viskas laisvai prieinama visiems: mokslui, verslui ir viešajam sektoriui.

Europos Sąjungos šalyse vykdomi projektai ir kuriami technologiniai sprendimai sparčiai įgyvendina semantinio saityno plėtrą nacionaliniu ir europiniu lygiu. Dėl straipsnio apimties apsiribosime tik svarbiausių iš jų apžvalga. Bendradarbiaudama su Sorbonos, Montrealio, Dortmundi ir kitais universitetais, Vokietijos įmonė TEMIS AG išvystė metaduomenų ištraukimo iš nestruktūrizuotos informacijos platforma *Luxid*¹⁵. Ši platforma yra universali. Ji gali būti taikoma nacionalinio saugumo, mokslinių tyrimų srityse, gali tenkinti tiek viešojo, tiek verslo sektorių poreikius. Bendradarbiaudama su Leipzigo universitetu, Vokietijos įmonė *Pattern Science AG* išvystė sintaksinės–semantinės analizės sprendimus¹⁶, kurie integruoti į taikomas programas *MyLOBSTER*, diegiamas stambiausioms Vokietijos žiniasklaidos, finansų bei verslo organizacijoms. Šefildo universitetas, bendradarbiaudamas su tarptautine korporacija *Ontotext AD*, sukūrė ir išvystė

¹⁴ Berners-Lee, T.; Hendlar, J.; Lassila, O. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American Magazine* [interaktyvus]. 2001 May [žiūrėta 2011.09.30]. <<http://www.scientificamerican.com/article.cfm?id=the-semantic-web>>.

¹⁵ LUXID [interaktyvus]. Heilderberg [žiūrėta 2011.09.30]. <<http://www.temis.com/index.php?id=201&selt=1>>.

¹⁶ Competitive Intelligence [interaktyvus]. Freigericht [žiūrėta 2011.09.30]. <<http://www.pattern-science.com/de/technologie/competitive-intelligence>>.

atviro kodo kalbos ir teksto inžinerijos platformą GATE (<http://gate.ac.uk/>), kurią *Ontotext AD* panaudojo žinių ir informacijos valdymo platformoje *KIM*, sprendžiančioje tekstų gavybos (angl. *Text Mining*) ir analizės, įvardytų esybių, įvykių identifikavimo, jų ryšių nustatymo, semantinio anotavimo ir semantinės paieškos uždavinius anglų kalbai¹⁷. Labai svarbus OpenCalais projektas (<http://www.opencalais.com/>), skirtas semantiniam anglų kalbos dokumentų anotavimui bei įvardytų esybių, įvykių ir faktų atpažinimui. Šio projekto sprendimai naudojami paieškos variklyje Interceder (<http://interceder.net/>), kuris leidžia ieškoti naujienų saityne anglų kalba (įskaitant socialinius tinklus) ir jas filtruoti pagal tematiką. Paminėtinas ir 6BP programos „Informacinės visuomenės technologijos“ imtyje vykdomas projektas *NeON* (<http://www.neon-project.org/>), kurio tikslas vystyti ontologijų kūrimo ir taikymo technologijas stambios apimties semantinėse aplikacijose. Svarbus 7BP projektas *ONTORULE* (<http://ontorule-project.eu/>), kurį vykdo Vienos, Paryžiaus ir Bolzano universitetai kartu su tokiomis stambiomis įmonėmis, kaip IBM, AUDI ir kt. Projekto tikslas - integruoti mokslo žinias ir technologijas, siekiant automatizuoti ontologijų kūrimą iš pasirinkamų šaltinių, įskaitant natūralios kalbos dokumentus. Europos Sąjungos iniciatyva ir lėšomis sukurta platforma *SEMIC EU* (*Semantic Interoperability Centre Europe* <http://www.semic.eu/>), skirta bendriems viešojo administravimo ir e. vyriausybės interoperabilumo ištekliams.

Palyginus lietuvių kalbos technologijų infrastruktūros situaciją su kitų ES šalių (Lenkijos, Vokietijos, Slovėnijos, Čekijos, Suomijos, Švedijos, Danijos, Latvijos, Estijos ir kt.) lygiu, galima konstatuoti, jog Lietuva nuo Europos Sąjungos atsilieka 7-9 metais. Lietuvių kalbos skaitmeniniai ištekliai ir įrankiai labai fragmentuoti, tarpusavyje nesuderinami, nėra numatytos bendros jų prieigos priemonės, jie saugomi skirtingose institucijose, skirtingais formatais, neužtikrinama pakankama prieiga per internetą ir sąveika „mašina-mašina“¹⁸. Galime teigti, kad semantinių technologijų taikymo užuomazgų Lietuvoje esama, bet jos apribotos siaurais specifiniais taikymais atskirose srityse. Tarp tokių semantinės analizės priemonių, ontologijų ir atskirų semantinių išteklių panaudojimo atvejų Lietuvoje galima paminėti *Infoplex.lt* teisinės informacijos paslaugas (<http://www.infoplex.lt/>), Lietuvos integralią bibliotekų informacinę sistemą *LIBIS* (<http://www.libis.lt/>), virtualią

¹⁷ KIM Platform [interaktyvus]. [žiūrėta 2011.09.30]. < <http://www.ontotext.com/kim> >.

¹⁸ Marcinkevičienė R., Vitkutė-Adžgauskienė, D. Developing the Human Language Technology Infrastructure in Lithuania. *Frontiers in Artificial Intelligence and Applications*. Vol. 219, *Human Language Technologies – the Baltic Perspective. Proceedings of the Fourth International Conference Baltic HLT 2010*. I. Skadina (ed.). Amsterdam: IOS Press, 2010: 3-10.

elektroninio paveldo sistemą VEPS (<http://www.epaveldas.lt>). Šie atskiri sprendimai remiasi unikaliais konkrečiai sričiai sukurtais semantiniiais ištekliiais ir įrankiais, juose praktiškai neišnaudojamos semantikos ir kalbos technologijų galimybės, dėl ko nukenčia jų kokybė. Lietuvių kalbai nėra parengta bendrųjų lingvistinės semantinės analizės įrankių, sudarančių galimybes trečiosioms šalims kurti semantinė analize grįstus sprendimus įvairiose taikymo srityse.

Lituanistinių skaitmeninių išteklių inventorizacija buvo atlikta Lietuvos Mokslo Tarybos finansuotame projekte „Lituanistinių skaitmeninių išteklių metaduomenų sistemos sukūrimas ir suderinimas su CLARIN“¹⁹, siekiant sudaryti prielaidas lietuvių kalbos išteklių informacinės infrastruktūros koncepcijai parengti. Paaiškėjo, kad nors apie 63% skaitmeninių išteklių turėtojų suteikia vienokią ar kitokią internetinę prieigą prie savo duomenų, tačiau dažniausiai apsiribojama pačia primityviausia informacijos paieškos galimybe. Iš pagrindinių skaitmeninių išteklių paminėtini Dabartinės lietuvių kalbos tekstynas (<http://donelaitis.vdu.lt>), Lietuvių mokslo kalbos tekstynas CorALit (<http://www.coralit.lt>), Dabartinės lietuvių kalbos žodynas (<http://www.lki.lt/dlkz/>). Esama situacija kliudo kurti patogias šiuolaikiškas elektronines paslaugas, grįstas semantinė analize. Be to, šie ištekliai kol kas nesiejami nei su nacionaline, nei su jokia tarptautine kalbos technologijų infrastruktūra. Tai rodo Lietuvos atskirtį nuo tarptautinės bendruomenės ir nesugebėjimą pasinaudoti jos patirtimi bei integracijos galimybėmis.

Europos strategijos forumas mokslinių tyrimų infrastruktūrų plėtrai (*ESFRI – European Strategy Forum for Research Infrastructures*,) yra suformulavęs ilgalaikę bendrą europinę mokslinių tyrimų infrastruktūrų (toliau – MTI) plėtros strategiją²⁰. Tarp humanitarinių ir socialinių mokslų srityje svarbiausiais pripažintų nacionalinių MTI projektų yra numatytas projektas „Elektroniniai lietuvių kalbos ištekliai (*E-lingua*)“ (Vytauto Didžiojo universitetas, Vilniaus universitetas, Šiaulių universitetas, Lietuvių kalbos institutas), tačiau nėra lėšų šio projekto vystymui. Nacionalinių MTI strategija remiasi esamų išteklių dokumentavimu bei unifikavimu, taip pat dalyvavimu atitinkamose transnacionalinėse MTI iniciatyvose, kurias kalbos išteklių ir technologijų srityje ryškiausiai atstovauja *CLARIN* (*Common Language Resources and Technology Infrastructure*) iniciatyva (<http://www.clarin.eu>). Pabrėžtina, kad suderinamumas su *CLARIN* rekomendacijomis ir

¹⁹ Lituanistiniai skaitmeniniai ištekliai [interaktyvus]. Kaunas: Vytauto Didžiojo universitetas [žiūrėta 2011.09.30]. <<http://sruoga.vdu.lt/lituanistiniai-skaitmeniai-istekliai>>.

²⁰ European Strategy Forum for Research Infrastructures [interaktyvus]. [žiūrėta 2011.09.30]. <http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri>.

standartais yra racionalus nacionalinės elektroninių lietuvių kalbos išteklių MTI kūrimo kelias.

Analizuojant kitų Europos šalių pavyzdžius, galima pastebėti, jog kalbos technologijų infrastruktūros plėtra koordinuojama valstybiniu mastu, derinami mokslo ir verslo poreikiai. Švedijoje 2007 m. startavo 7 institucijų projektas „Švedų kalbos technologijų infrastruktūra“, danai suformavo *CLARIN* konsorciumą, kuriame 4 universitetai ir 4 kultūros įstaigos bendrai kuria nacionalinę su *CLARIN* suderinamą kalbos išteklių infrastruktūrą. Analogišką konsorciumą (*FIN-CLARIN*) turi suomiai²¹. Estai kalbos technologijų klausimus yra įtraukę į ilgalaikę estų kalbos vystymo strategiją. Visos šios iniciatyvos numato kuriamos kalbos technologijų infrastruktūros suderinimą su tarptautinėmis kalbos technologijų kūrimo iniciatyvomis.

Lietuvoje tikru proveržiu, siekiant suderinti mokslo, viešojo ir privataus sektorių interesus semantinių ir kalbos technologijų srityje, laikytina 2011 m. paskelbta Ekonomikos augimo veiksmų programos 3 prioriteto „Informacinė visuomenė visiems“ įgyvendinimo priemonė „Lietuvių kalba informacinėje visuomenėje“. Vienos iš veiklų, kurios bus finansuojamos, numato esamo lietuvių kalbos tekstyno papildymą, lietuvių kalbos sintaksinės-semantinės analizės priemonių ir jų pritaikymų kūrimą, lietuviškų interneto svetainių analizės sistemos ir jos pritaikymų kūrimą²². Konkursą laimėjusio projekto įgyvendintojai privalės sukurti informacinę sistemą, saugančią lietuvių kalbos ir analizės priemones ir rezultatus, pritaikytą analizuoti lietuviško turinio interneto svetainių turinį, atlikti pagal jį paiešką, palyginti tarpusavyje svetainėse esančius tekstus ir turinį, pateikti analizės rezultatus vartotojui priimtina forma²³. Kadangi paraiškos pagal šią priemonę dar vertinamos, todėl projektų turinys neskelbtinas ir nekomentuotinas, bet labai tikėtina, kad šį kartą pagaliau pavyks patenkinti tiek mokslo, tiek privataus, tiek viešojo sektorių interesus (įskaitant elektroninės erdvės kontrolę).

Apibendrinant darytina išvada, kad semantinių technologijų kūrimo ir taikymo užuomazgų Lietuvoje esama, bet jos apribotos siaurais specifiniais taikymais atskirose srityse, dažniausiai – kalbotyros srityje. Esami lietuvių kalbos semantiniai ištekliai nepakankami, o bazinės technologijos nepakankamai išvystytos, todėl lietuvių rašytinės kalbos semantinės analizės paslaugų ir technologinių sprendimų pasiūla netenkina paklausos, neišnaudojamos

²¹ FIN-CLARIN [interaktyvus]. University of Helsinki. Department of Modern Languages [žiūrėta 2011.09.30]. <<http://www.ling.helsinki.fi/finclarin/intro.html>>.

²² Projektų, vykdomų pagal Ekonomikos augimo veiksmų programos 3 prioriteto „INFORMACINĖ VISUOMENĖ VISIEMS“ įgyvendinimo priemonę Nr. VP2-3.1-IVPK-12-K „Lietuvių kalba informacinėje visuomenėje“, finansavimo sąlygų aprašas. *Valstybės žinios*. 2011, Nr.: 30-1431: punktas 4.2.

²³ *Ibid.*, punktas 94.3.

nacionalinės ir tarptautinės integracijos galimybės, iškyla grėsmė rašytinės lietuvių kalbos išsaugojimui internete. Ši situacija susidarė dėl trūkumo bendradarbiavimo tarp privataus ir viešojo sektoriaus, bei akademinės bendruomenės, kuris labai efektyviai vyksta kitose Europos Sąjungos valstybėse.

SMURTO KONTROLĖS LIETUVIŠKAME SAITYNE SAŠAJOS SU SEMANTINIŲ IR KALBOS TECHNOLOGIJŲ PLĖTRA

Siekiant išryškinti smurto kontrolės lietuviškame saityne sąryšį su semantinių ir kalbų technologijų plėtra, reikia atsakyti į tris klausimus: ar elektroninė erdvė yra viešoji erdvė, ar reikia vykdyti elektroninės erdvės kontrolę ir kuo šioje veikloje teisėtvarkos organams gali būti naudingos semantinės ir kalbos technologijos?

Į pirmą klausimą vienareikšmio atsakymo nėra. Iš vienos pusės, reikia pripažinti, kad susidariusi teisinė spraga, nes dabartiniai teisės aktai elektroninės erdvės neapibrėžia kaip viešosios erdvės. Iš kitos pusės, viešų ir el. valdžios paslaugų teikimas bent dalį elektroninės erdvės paverčia viešąja erdve. Be to, Lietuvos pilietis, besilankantis socialiniuose tinklalapiuose ar kituose elektroninės erdvės ištekliuose, nori jaustis saugus. Dėl šios priežasties susidaro paradoksali situacija. Kai pilietis ar jo vaikas saityne patiria smurtą ar patyčias, jis retoriškai užduoda klausimą – kur žiūri policija? Bet, remiantis dabartiniais teisės aktais ir neturint priemonių, policijos pareigūnams belieka vykdyti prevencinę veiklą arba pagal gautą faktologinę medžiagą tirti jau įvykdytą konkretų elektroninį nusikaltimą. Darytina išvada, kad viešąja elektronine erdve laikytina ta saityno erdvės dalis, prie kurios prieigos neriboja slaptažodis arba kita apsaugos priemonė.

Atsakant į antrą klausimą, labai reikšmingas Nacionalinės nusikaltimų prevencijos ir kontrolės programos²⁴ (toliau - Programa) reglamentavimas ir teiginiai, bei aukščiau minėta Ekonomikos augimo veiksmų programos 3 prioriteto „Informacinė visuomenė visiems“ įgyvendinimo priemonė „Lietuvių kalba informacinėje visuomenėje“.

Programoje apibrėžta, kad nusikaltimų kontrolė – poveikio nusikalstamumui priemonė, kuria siekiama mažinti nusikalstamumą ir neleisti viršyti socialiai priimtino lygmens aktyviais teisėsaugos institucijų veiksmais ir teisinio baudžiamojo poveikio priemonėmis, taip pat aktyviomis administracinėmis, ekonominėmis, socialinėmis, kitokio pobūdžio priemonėmis. Programos 6 punktą konstatuoja, kad asmens ir visuomenės saugumas yra būtina pagrindinio

²⁴ Nacionalinė nusikaltimų prevencijos ir kontrolės programa prevencijos programa. *Valstybės žinios*. 2003, Nr.: 32 - 1318.

valstybės politikos tikslo - konstitucinių žmogaus teisių ir laisvių užtikrinimo – įgyvendinimo sąlyga. Siekdama užtikrinti kiekvieno individo, o kartu ir visuomenės saugumą, valstybė vykdo nusikaltimų prevenciją ir kontrolę. Labai svarbi Programos 8 punkte išryškinta sąlyga, kad nusikaltimų prevencija ir kontrolė gali būti sėkminga, jeigu joje aktyviai dalyvauja visi socialiniai subjektai – teisėsaugos institucijos, kitos viešojo administravimo institucijos, savivaldybių institucijos, verslo bendruomenė, visuomeninės organizacijos ir individualūs asmenys. Ši veikla turi būti koordinuota ir remtis mokslu bei praktika pagrįsta valstybės politinių partijų patvirtinta bendra politika. Aštuonioliktas punktas pripažįsta, kad nusikaltimų prevencijos ir kontrolės organizavimas dar nėra pakankamai nuoseklus ir sistemingas, neatitinka nusikaltimų prevencijos bei kontrolės principų. Pagrindinės to priežastys: 1) tikslinių programų vykdytojų interesai ir tikslai labiau orientuojami į planuotų priemonių įvykdymą nei į nuolatinę veiklą nusikaltimų prevencijos ir kontrolės srityje. Šių programų įgyvendinimas ir finansavimas baigiasi jose nustatytu terminu, nors nusikaltimų prevencijos ir kontrolės organizavimo poreikis yra nuolatinis. Dažnai priemonėms įgyvendinti ar naujoms priemonėms kurti neskiriama pakankamai lėšų. Dvidešimtame punkte atskleidžiami nusikaltimų prevencijos ir kontrolės trūkumai. Vienu iš pagrindinių yra tai, kad nusikaltimų prevencijos ir kontrolės tikslams nėra panaudojamos kitų socialinių, ekonominių, urbanistinių ir kitokių programų galimybės. Tokios programos šalyje nėra net analizuojamos ir vertinamos kriminogeninių veiksnių šalinimo, neutralizavimo ar kompensavimo požiūriais. Atkreiptinas dėmesys į Programos 53 punkte pateiktą prioritетinių nusikaltimų prevencijos ir kontrolės kryptių dispoziciją. Iš vienuolikos prioritетinių kryptių pirmoje vietoje yra „smurtinių nusikaltimų prevencija ir kontrolė“, trečioje – „nusikaltimų, daromų viešosiose vietose, prevencija ir kontrolė“, vienuoliktoje (paskutinėje) – „vaikų ir jaunimo nusikalstamumo prevencija ir kontrolė“. 24 punkte išdėstyti nusikaltimų prevencijos ir kontrolės veiklos efektyvumo principai, iš kurių čia paminėtini aprūpinimo užtikrinimo ir mokslinio pagrįstumo principai. Efektyvus poveikis nusikalstamumui įmanomas tik esant tinkamam politiniam, organizaciniam, finansiniam, materialiniam, informaciniam, moksliniam, profesiniam užtikrinimui ir palaikymui, o nusikaltimų prevencijos ir kontrolės planai turi būti pagrįsti moksliskai, jie turi būti vykdomi tik esant pakankamai racionalių mokslinių argumentų, rodančių tikėtiną šių planų veiksmingumą. Programoje apie elektroninę erdvę tiesiogiai nekalbama. Tą patį pastebime ir Nacionalinės nusikaltimų prevencijos ir kontrolės 2010-2012

įgyvendinimo priemonių plane, kurio tik dešimtas punktą numato sukčiavimo elektroninėje erdvėje mokslinio įvertinimo priemonę²⁵.

Ekonomikos augimo veiksmų programos 3 prioriteto „Informacinė visuomenė visiems“ įgyvendinimo priemonė „Lietuvių kalba informacinėje visuomenėje“ sudarytų sąlygas bendradarbiavimui tarp privataus ir viešojo sektoriaus, bei akademinės bendruomenės, o taip priemonių elektroninės erdvės kontrolei sukūrimui ir įgyvendinimui. Tačiau priemonės finansavimo sąlygų aprašo 5 punkte įvedamas ribojimas „Veiklų tikslas – taikant šiuolaikines informacines ir ryšių technologijas, kurti sprendimus, padedančius išsaugoti lietuvių kalbą visose viešojo valstybės gyvenimo srityse <...>“, nors to paties dokumento 23 punkte nurodoma, kad projektai, įgyvendinami pagal šią Priemonę, turi atitikti Ekonomikos augimo veiksmų programos 3 prioriteto „Informacinė visuomenė visiems“ 1 uždavinį „plėtoti elektroninius sprendimus, didinant viešojo sektoriaus institucijų efektyvumą <...>“. Darytina išvada, kad komunikacija tarp valstybinių programų rengėjų ir viešojo sektoriaus suinteresuotų institucijų buvo nepakankama, todėl labai abejotina, kad projektų kūrėjai ir vertintojai atsižvelgs į viešojo saugumo ir viešosios tvarkos užtikrinimo interesus.

Liko atsakyti į trečią klausimą – kuo šioje veikloje teisėtvarkos organams gali būti naudingos semantinės ir kalbos technologijos? Nors internete vis daugiau informacijos perduodamas garsu ir vaizdu, bet tekstas vis dar lieka pagrindine informacijos perdavimo forma. Todėl didelius kiekius tekstinės informacijos ar kitokių nestruktūrizuotų duomenų analizuojančios teisės saugos bei šalies saugumo užtikrinimo ir atskirų sričių politikos įgyvendinimo institucijos (Policijos departamentas Užsienio reikalų ministerija, Krašto apsaugos ministerija, Valstybės saugumo departamentas, Specialiųjų tyrimų tarnyba ir kt.) susiduria su problema vykdydamos informacijos lietuvių kalba paiešką ir analizę – nėra veiksmingų prasminės (semantinės) teksto paieškos ir analizės priemonių (informacinių taikomųjų sistemų), pritaikytų dirbti su lietuvių kalbos tekstais ir duomenimis. Pabrėžtina, kad minėtų institucijų vartotojai didelio vartotojų srauto nesukuria, todėl į jų poreikius dažniausiai neatsižvelgiama, bet jų darbo rezultatas daro tiesioginę teigiamą įtaką kitoms valstybės veiklos sritims. Ši tikslinė grupė susiduria su lietuvių kalbai pritaikytų saityno turinio semantinės paieškos ir analizės priemonių trūkumu, lyginat su tuo, kas Pasauliniame tinkle jau sukurta kitoms kalboms. Kalbant apie konkrečias priemones, būtina sukurti specializuotas ontologijas ir saugyklą, specializuotus tekstynus, esybių, informacijos ir ryšių sintaksinės –

²⁵ Nacionalinės nusikaltimų prevencijos ir kontrolės programos 2010 - 2012 metų įgyvendinimo priemonių planas. *Valstybės žinios*. 2010, Nr.: 31-1433.

semantinės paieškos ir analizės viešuose tekstuose lietuvių kalba priemonės, faktų, įvykių, procesų ar veikimo stebėsenos (informacijos apie faktinius ar galimus įvykius, procesus ar veikimą, kurie gali kelti grėsmę nacionaliniam saugumui) paieška pagal sukurtus šablonus (grėsmių scenarijus) priemonės; grupių nuotaikų, jų kitimo tendencijų įvertinimo esybės atžvilgiu priemonės²⁶. Sukurtos ir įdiegtos semantinių ir kalbos technologijų naujovės gali užtikrinti teisėsaugos bei šalies saugumą užtikrinančių institucijų darbo kokybę, įgyvendinant viešąjį saugumą, kas turėtų teigiamą poveikį visai valstybei.

Apibendrinant galima teigti, kad elektroninės erdvės kontrolė, siekiant vykdyti smurto lietuviškame saityne efektyvią kontrolę ir užtikrinti visuomenės saugumą, dabartiniuose teisės aktuose įžvelgiama tik netiesiogiai. Probleminė situacija dėl elektroninės erdvės kontrolės ir semantinių bei kalbos technologijų panaudojimo šioje srityje susidarė dėl vėluojančių valstybės strateginių programų ir priemonių, bendros kalbos technologijų išteklių ir įrankių vystymo koncepcijos nebuvimo ir nepakankamų valstybės investicijų į semantinio pasaulinio tinklo technologijas, riboto bendradarbiavimo tarp privataus ir viešojo sektoriaus, bei akademinės bendruomenės. Kaip rodo pasaulinė praktika, šio tipo investicijos tiesioginės finansinės naudos neduoda, nesukuria didelio vartotojų srauto, todėl verslui jos nepatrauklios. Be to, universitetai arba privatus sektorius nepajėgūs įgyvendinti tokios apimties uždavinių be valstybės paramos.

IŠVADOS

Smurtas elektroninėje erdvėje jau tapo tarptautine problema, o elektroninė erdvę galima prilyginti „laukiniams Vakarams“, kur žmogui atsiveria neribotos galimybės, bet kur nevaržomai plinta smurtas ir prievarta, nes niekadėjai nesulaukia pelnytos bausmės. Intensyvi interneto ir informacinių technologijų sklaida gyventojų tarpe leidžia prognozuoti, kad nesant efektyvių poveikio priemonių, smurto atvejų elektroninėje erdvėje tik daugės. Auka dažnai nutyli apie patirtą elektroninį smurtą ir, dažniausiai, negali pati apsiginti. Tai rodo, kad, siekiant užtikrinti piliečių saugumą internete, efektyvių prevencijos priemonių jau nepakanka. Būtinės veiksmingos elektroninės erdvės kontrolės priemonės.

Nors nemažai informacijos internete jau perduodama vaizdu ir garsu, bet pagrindine forma iki šiol lieka tekstas. Informacijos kiekis sparčiai didėja, esami skaičiavimo metodai jau

²⁶ Straipsnio autorius 2011 m. birželio mėn. atliko specialiųjų tarnybų vadovų ir specialistų apklausą, kurios tikslas buvo išsiaiškinti semantinių ir kalbos technologijų sprendimų poreikius ir taikymo galimybes specialiųjų tarnybų veikloje. Pagal apklausos sąlygas detalūs apklausos rezultatai konfidencialūs, publikuoti tik apibendrinti sprendimų poreikiai.

nesugeba jo apdoroti. Todėl reikalingos kokybiškai naujos, semantinės ir kalbų technologijos, kurios mokėtų analizuoti ir apibendrinti informaciją pagal jos prasmę. Semantinių technologijų kūrimo ir taikymo užuomazgų Lietuvoje esama, bet jos apribotos siaurais specifiniais taikymais atskirose srityse, dažniausiai - kalbotyros srityje. Šioje srityje Lietuva nuo Europos Sąjungos atsilieka 7-9 metais. Esami lietuvių kalbos semantiniai ištekliai nepakankami, o bazinės technologijos nepakankamai išvystytos, todėl lietuvių rašytinės kalbos semantinės analizės paslaugų ir technologinių sprendimų pasiūla netenkina paklauskos, neišnaudojamos nacionalinės ir tarptautinės integracijos galimybės, išskyla grėsmė rašytinės lietuvių kalbos išsaugojimui internete. Ši situacija susidarė dėl trūkumo bendradarbiavimo tarp privataus ir viešojo sektoriaus, bei akademinės bendruomenės, kuris labai efektyviai vyksta kitose Europos Sąjungos valstybėse.

Didelius kiekius tekstinės informacijos ar kitokių nestruktūrizuotų duomenų analizuojančios teisėsaugos bei šalies saugumo užtikrinimo ir atskirų sričių politikos įgyvendinimo institucijos susiduria su problema vykdydamos informacijos lietuvių kalba paiešką ir analizę – nėra veiksmingų prasminės (semantinės) teksto paieškos ir analizės priemonių, pritaikytų dirbti su lietuvių kalbos tekstais ir duomenimis. Sukurtos ir įdiegtos semantinių ir kalbos technologijų naujovės gali užtikrinti teisėsaugos bei viešąjį saugumą užtikrinančių institucijų darbo kokybę, įgyvendinant viešąjį saugumą elektroninėje erdvėje, kas turėtų teigiamą poveikį visai valstybei. Tačiau tai padaryti kliudo vėluojančios valstybės strateginės programos ir priemonės, bendros kalbos technologijų išteklių ir įrankių vystymo koncepcijos nebuvimas, nepakankamos valstybės investicijos į semantinio pasaulinio tinklo technologijas, ribotas bendradarbiavimas tarp privataus ir viešojo sektoriaus, bei akademinės bendruomenės.

LITERATŪRA

1. Nacionalinė nusikaltimų prevencijos ir kontrolės programa prevencijos programa. *Valstybės žinios*. 2003, Nr.: 32 - 1318.
2. Nacionalinės nusikaltimų prevencijos ir kontrolės programos 2010 - 2012 metų įgyvendinimo priemonių planas. *Valstybės žinios*. 2010, Nr.: 31-1433.
3. Projektų, vykdomų pagal Ekonomikos augimo veiksmų programos 3 prioriteto „INFORMACINĖ VISUOMENĖ VISIEMS“ įgyvendinimo priemonę Nr. VP2-3.1-IVPK-12-K „Lietuvių kalba informacinėje visuomenėje“, finansavimo sąlygų aprašas. *Valstybės žinios*. 2011, Nr.: 30-1431.
4. Europe's Digital Competitiveness Report [žiūrėta 2011.09.30]. <http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=667>.

5. Europe's Digital Competitiveness Report - ICT Country Profiles [žiūrėta 2011.09.30]. <http://ec.europa.eu/information_society/newsroom/cf/document.cfm?action=display&doc_id=668>.
6. European Strategy Forum for Research Infrastructures [interaktyvus]. [žiūrėta 2011.09.30]. <http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri>.
7. Berners-Lee, T.; Hendler, J.; Lassila, O. The Semantic Web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American Magazine* [interaktyvus]. 2001 May [žiūrėta 2011.09.30]. <<http://www.scientificamerican.com/article.cfm?id=the-semantic-web>>.
8. Cheung, A. S.Y. A Study of Cyber-Violence and Internet Service Providers' liability: Lessons from China. *Pacific Rim Law & Policy Journal*. 2009, 18(2): 323-346.
9. Competitive Intelligence [interaktyvus]. Freigericht [žiūrėta 2011.09.30]. <<http://www.pattern-science.com/de/technologie/competitive-intelligence>>.
10. Elektroninių patyčių paplitimas Klaipėdos apskrities vyresniųjų moksleivių tarpe (tyrimo duomenys) [interaktyvus]. Klaipėdos universiteto Psichologijos katedra, 2009 [žiūrėta 2011.09.30]. <http://www.ku.lt/pf/psichologija/doc/Tyrimo_duomenys_1.pdf>.
11. FIN-CLARIN [interaktyvus]. University of Helsinki. Department of Modern Languages [žiūrėta 2011.09.30]. <<http://www.ling.helsinki.fi/finclarin/intro.html>>
12. Informacinės Visuomenės Plėtros Komiteto prie Susisiekimo ministerijos statistika [interaktyvi]. Vilnius: 2011 [žiūrėta 2011.09.30]. <<http://www.ivpk.lt/lt/ltvm/statistika>>.
13. Interneto, kompiuterinių žaidimų įtaka smurtui prieš vaikus [interaktyvus]. Vilnius: Socialinės apsaugos ir darbo ministerija, UAB „Factus Dominus“, 2005 [žiūrėta 2011.09.30]. <<http://www.draugiskasinternetas.lt/lt/adult/news?id=2003>>.
14. Kane, J.; Portin, P., *Smurtas ir technologijos* [interaktyvus]. Briuselis, 2008 [žiūrėta 2011.09.30] <http://ec.europa.eu/justice_home/daphnetoolkit/files/others/booklets/07_daphne_booklet_7_lt.pdf>.
15. KIM Platform [interaktyvus]. [žiūrėta 2011.09.30]. <<http://www.ontotext.com/kim>>.
16. Lietuvos Respublikos nacionalinis elektroninių ryšių tinklų ir informacijos saugumo incidentų tyrimo padalinys (CERT-LT): metinė ir ketvirtinė incidentų statistika [interaktyvus]. Vilnius, 2011 [žiūrėta 2011.09.30]. <<https://www.cert.lt/statistika.html>>.
17. Lituanistiniai skaitmeniniai ištekliai [interaktyvus]. Kaunas: Vytauto Didžiojo universitetas [žiūrėta 2011.09.30]. <<http://suoga.vdu.lt/lituanistiniai-skaitmeniai-istekliai>>.
18. LUXID [interaktyvus]. Heilderberg [žiūrėta 2011.09.30]. <<http://www.temis.com/index.php?id=201&self=1>>.
19. Marcinkevicienė R., Vitkutė-Adžgauskienė, D. Developing the Human Language Technology Infrastructure in Lithuania. *Frontiers in Artificial Intelligence and Applications. Vol. 219, Human Language Technologies – the Baltic Perspective. Proceedings of the Fourth International Conference Baltic HLT 2010*. I.Skadina (ed.). Amsterdam: IOS Press, 2010: 3-10.
20. Morriss, A. P. The Wild West Meets Cyberspace. *The Freeman: Ideas on Liberty* [interaktyvus]. 1998, 48(7) [žiūrėta 2011.09.30]. <<http://www.thefreemanonline.org/featured/the-wild-west-meets-cyberspace/print/>>.
21. Pilkauskaitė Valickienė R.; Raižienė, S.; Žukauskienė, R. Elektroninių Patyčių Paplitimas tarp Klaipėdos Apskrities Vyresniųjų Klasių Moksleivių. *Socialinis darbas*. 2009, 8(2): 114-121.
22. Projektas FIN-CLARIN [interaktyvus]. University of Helsinki. Department of Modern Languages: 2011. [žiūrėta 2011.09.30]. <<http://www.ling.helsinki.fi/finclarin/intro.html>>.
23. Ruškus, J.; Žvirdauskas, D.; Kačenauskaitė, V.; Žvirdauskienė R. *Paauglių elgsena elektroninėje erdvėje*. Švietimo informacinių technologijų centro 2009 m. atliktas tyrimas [žiūrėta 2011.09.30]. <http://www.draugiskasinternetas.lt/repository/dokumentai/11_03_24/tyrimas%20elgsena%20elektroneje%20aplinkoje.pdf>

-
24. Smith P.K.; Morita, Y. *The Nature of School Bullying. A cross-national perspective*. In P.K. Smith et al (eds). London and New York: Routledge, 1999.

VIOLENCE CONTROL IN LITHUANIAN WORLD WIDE WEB RELATION WITH THE DEVELOPMENT OF SEMANTIC AND NATURAL LANGUAGE TECHNOLOGIES

Darius Amilevičius*
Mykolas Romeris University

Summary

Violence in cyberspace has become an international problem. The World Wide Web is comparable to "the wild West," where a person has big opportunities, but they are spreading violence, because wicked men fails to receive due punishment. Intensive Internet and information technologies broadening across the population allows us to predict that, in the absence of the effective control, in future violence will increase in the Cyberspace. Victim is often silent about their experiences of cyber-violence, and often cannot defend itself. This suggests that, in order to ensure the public safety on the Internet, effective prevention measures are no longer sufficient. It is necessary to proceed with effective cyber controls.

Actually amount of video and audio information is increasing, but the main form still remains the text. The general amount of information is increasing rapidly, the existing methods of calculation is not able to process it. It is therefore necessary for qualitatively new semantic and linguistic technologies solutions to analyze and summarize the information according to its meaning. Eventually the germs of Semantic technologies in Lithuania exist, but it's limited by narrow specific applications in different fields, mostly - in linguistics researches. In this area Lithuania lags 7-9 years compared to the states of European Union. The current Lithuanian language and semantic resources are inadequate, the basic technologies are not sufficiently developed, therefore, , the technological solutions and services of the semantic analysis of Lithuanian written language does not meet the demand, under-used national and international integration opportunities, there is a threat to the preservation of the written Lithuanian language on the Internet. This situation is due to the lack of cooperation between the private and public sector, and academic researchers, which is very effective in other European Union countries.

The law enforcement and public security authorities, which must proceed with analysis of large amounts of textual information or other unstructured data, are faced with the problem of effective search and analysis of information in Lithuanian language - there is no effective semantic search and text analysis tools tailored to work with Lithuanian language text and data. Developed and implemented semantic language technologies and innovations can to ensure to law enforcement and public security authorities the quality of the implementation of public security in cyberspace who would have a positive impact on the entire state. However, it is hindered by delays in state strategic framework and tools, by the absence of the concept of language technologies and tools development, by the lack of public investment in a global network of semantic technologies, by the limited cooperation between the private and public sector, and researchers.

Keywords: cyber-violence control, semantic technology, natural language technology, security in cyberspace.

Darius Amilevičius*, Mykolas Romeris universiteto Viešojo saugumo fakulteto Humanitarinių mokslų katedros docentas. Mokslinių tyrimų kryptys: politinė ir juridinė retorika, semantinės ir kalbų technologijos, marketinginė komunikacija, žmogaus teisės.

Darius Amilevičius*, Mykolas Romeris University, Faculty of Public security, Department of Humanities, assoc. professor. Research interests: political and juridical rhetoric, semantic and natural language technologies, marketing communication, human rights.