



Olga Zervina

**AUTOMATIZĒTĀS VĒRTĪBU NOTEIKŠANAS
METODOLOĢIJA UZ E-TEKSTU ANALĪZES PAMATA
GAISA TRANSPORTA NOZARĒ**

**A METHODOLOGY OF AUTOMATED IDENTIFICATION
OF VALUES IN THE AIR TRANSPORTATION DOMAIN ON
THE BASE OF E-TEXTS ANALYSIS**

Promocijas darba kopsavilkums

zinātniskā doktora grāda iegūšanai
būvniecības un transporta inženierzinātnēs
Apakšnozare – telemātika un loģistika

Summary of the promotion work

to obtain the scientific degree Doctor of Science in Engineering
Scientific area “Civil Engineering and Transport”
Scientific subarea “Telematics and Logistics”

Zinātniskais vadītājs:

Stukalina Yulia, Dr.sc.admin.
Pavlyuk Dmitry, Dr.sc.ing., Dr.oec
Neil Rubens, PhD

RĪGA – 2022

UDK

Transporta un sakaru institūts
Transport and Telecommunication Institute

O. Zervina

Automatizētās vērtību noteikšanas metodoloģija uz e-tekstu analīzes pamata gaisa transporta nozarē: Promocijas darba kopsavilkums. Rīga, Transporta un sakaru institūts, 2022.

A methodology of automated identification of values in the air transportation domain on the base of e-texts analysis: Summary of the promotion work. Riga: Transport and Telecommunication Institute, 2022.

Iesniegtajā materiālā saglabāts autora stils un noformējums.
The submitted material retains the author's style and design.

ISBN 978-9934-622-00-7

© Olga Zervina, 2022

© Transporta un sakaru institūts, 2022

**PROMOCIJAS DARBS IESNIEGTS
TRANSPORTA UN SAKARU INSTITŪTĀ
ZINĀTNISKĀ DOKTORA GRĀDA IEGŪŠANAI
BUVNICĪBAS UN TRANSPORTA INŽENIERZINĀTNĒS**

OFICIĀLIE RECENZENTI:

Dr.sc.ing. Irina Jackiva
Profesore, Transporta un sakaru institūts, Rīga, Latvija

Dr. Eftihia Nathanail
Professor of Transportation Engineering, University of Thessaly,
Griekija

Dr. Vera Sheinman
NLP Software Engineer, Apple Inc., Cambridge, Lielbritānija

Promocijas darba aizstāvēšana notiks 2022.gada 21. decembrī, plkst. 15:00
Transporta un sakaru institūta promocijas padomē pēc adreses: Lomonosova iela 1,
130. aud., Rīga, Latvija, tālr. (+371) 67100594, fakss: (+371) 67100535.

APSTIPRINĀJUMS

Es apstiprinu, ka esmu izstrādājusi doto promocijas darbu, kas iesniegts izskatīšanai
Transporta un sakaru institūta zinātniskā doktora grāda iegūšanai būvniecības un
transporta inženierzinātnēs. Šis promocijas darbs iepriekš nav iesniegts izskatīšanai
citās promocijas padomēs.

O. Zervina

Promocijas darbs ir uzrakstīts angļu valodā, sastāv no 5 nodaļām, iekļauj 39 attēlus,
24 tabulas un 14 pielikumus, 194 lappuses. Izmantotās literatūras sarakstā ir 301
avoti.

SATURS

ANOTĀCIJA	5
1. PĒTĪJUMA AKTUALITĀTE UN MOTIVĀCIJA	6
2. PĒTĪJUMA MĒRĶI UN UZDEVUMI	7
3. PĒTĪJUMA MĒRĶIS, OBJEKTS UN PRIEKŠMETS	7
4. PĒTNIECĪBAS DARBĪBAS JOMA	7
5. PIEEJA.....	9
6. PĒTĪJUMA METODOLOĢIJA UN METODEDES.....	9
7. PĒTĪJUMĀ ZINĀTNISKĀ NOVITĀTE.....	9
8. PĒTĪJUMA TEORĒTISKĀ UN PRAKTISKĀ VĒRTĪBA	10
9. PĒTĪJUMA APPROBĒŠANA	10
10. PROMOCIJAS DARBA STRUKTŪRA.....	11
11. AIZSTĀVĒŠANAI IZVIRZĪTĀS TĒZES	11
12. PROMOCIJAS DARBA NODAĻU KOPSAVILKUMS.....	12
12.1. Vērtības kā vērtības piedāvājuma daļa gaisa pārvadājumos	12
12.2. Metodoloģiskās pieejas: Pašreizējais stāvoklis/State of art.....	14
12.3. Metodoloģija un pētījumu dizains	17
12.4. Galvenie atklājumi un izpētes rezultātu īstenošana.....	22
12.5. Pētniecības rezultātu diskusija.....	27
Secinājumi un ieteikumi	28
Autora publikāciju bibliogrāfiskais saraksts.....	56
Literatūras saraksts	57

ANOTĀCIJA

Promocijas darbu “Automatizētās vērtību noteikšanas metodoloģija uz e-tekstu analīzes pamata gaisa transporta nozarē” izstrādājusi Olga Zervina Dr.sc.admin, profesore Jūlijas Stukaļinas, Dr.sc.ing., profesora Dmitrija Pavļuka un Dr. Nīla Rubensa vadībā.

Kā pētījumu objekts šajā darbā tiek noteikts vērtības piedāvājums gaisa transporta jomā. Darbs ir vērsts uz datu balstīta lēmumu pieņemšanas atbalsta modeļa izstrādi, nosakot instrumentus un procedūras, ko lēmumu pieņēmēji varētu izmantot, lai noteiktu jaunas vērtības kā daļu no vērtības izvirzīšanas.

Prasība pēc jaunām vērtībām un vērtību skaita palielināšanas gaisa transportā spilgti parādīta 2018. gadā Starptautiskās gaisa transporta asociācijas (IATA) nozares lietu komitejas (IAC) ziņojumā “Avioindustrijas nākotne 2035”. Dokumenta 12. ieteikumā ir skaidri norādīts, ka, “tā kā jaunās tehnoloģijas un vērtību maiņas maina veidu, kā un kāpēc cilvēki strādā, nozarei būs jāiegulda prasmēs nākamajiem aviācijas vadītājiem un darbiniekiem un jāinformē par ieguvumiem, ko sniedz darbs šajā nozarē”.

Arguments par plašāku aviācijas vērtību ir atspoguļots IATA Rūpniecības lietu komitejas ziņojumā kā Lielo datu, automatizācijas un interneta jautājumu tendenču kombinācija. Paredzams, ka šīs izmaiņas radīs jaunas iespējas, mainot cilvēku un ierīču pārraudzību un mārketingu reāllaikā. Gaisa transporta uzņēmumiem ir izdevīgi saņemt atbildību par automatizācijas sasniegumiem, jauniem transporta veidiem un patērētāju attieksmi.

Šis pētījums atbalsta koncepciju, ka vērtības datu kontrolei būs priekšrocības salīdzinājumā ar esošajiem konkurentiem jaunu nišu izstrādē. Pētījuma novitāte un mērķis slēpjas vismodernākajā metodoloģijā, kas ļauj teksta ieguves metodes kā daļu no Dabiskās valodas apstrādes izmantot daudzās jaunās lapās.

Vērtību noteikšana ir kompleksa un sarežģīta problēma. Autors izmanto NLP pieeju, lai apzinātu šai problēmai labi piemērotas metodes un iezīmes. Tiek izmantota uz datiem balstīta tehnika, t.i., autore vispirms izveido datu kopu un pēc tam analizē to, izmantojot skaitļošanas lingvistiskās metodes, lai noteiktu, kuras funkcijas un metodes darbojas vislabāk. Autore izmanto arī korpuslingvistisko pieeju, t.i., uzticama valodas analīze ir vairāk iespējama ar korpusiem, kas savākti laukā tās dabiskajā kontekstā.

Galvenie šī zinātniskā darba rezultāti ir prezentēti 14 starptautiskās zinātniskās konferencēs un publicēti 12 zinātniskos referātos. Zinātniskais darbs sastāv no 5 nodaļām un ietver 194 lappuses, 39 attēlus, 24 tabulas pamattekstā, 14 pielikumus un 301 bibliogrāfiskās atsauces.

1. PĒTĪJUMA AKTUALITĀTE UN MOTIVĀCIJA

Vērtības izveide ir viens no jebkuras biznesa organizācijas primārajiem mērķiem (Prasada et al, 2021). Uzņēmuma stratēģija sākas ar tā vērtības piedāvājuma izstrādi (Lanning and Michaels, 1988) nosaucot vienkāršu paziņojumu, kurā apkopots, kāpēc klientam vajadzētu izvēlēties tā produktu vai pakalpojumu: mūsu pieejamās lidmašīnas biļetes, mūsu ērtā biznesa klases atpūtas telpa. Ja uzņēmuma vērtības piedāvājumu apņemas veidot vairāki cilvēki, viņu rezultāti var atšķirties gan fokusa, gan temata ziņā. Un jautājums ir, kurš no šiem vērtību piedāvājumiem ir pareizs? Varbūt tie visi neatspoguļo tirgus vērtības? Lai to noskaidrotu, uzņēmumiem ir jāveic dārgi un ilgtermiņa mārketinga pētījumi, kas ietver diezgan daudz cilvēku kļūdu. Situācijas uzlabošanai ir jāatrod objektīvāka pieeja, iespējams, ar tehnisko atbalstu, kas ir diezgan iespējams ar mūsdienu pasaules potenciālu (Ebel et al., 2016).

Vērtības noteikšanas automatizācija noteikti novestu pie vērtību skaita palielināšanas. Turklāt Kaplans un Nortons (2008) iesaka uzņēmuma stratēģiju balstīt uz “diferencētu klientu vērtības piedāvājumu”. Jaunākie pētījumi parāda milzīgo atšķirību starp uzņēmumu darbības rezultātiem atkarībā no klientiem piedāvāto vērtību skaita. Kad pircēji novērtē pakalpojumu vai produktu, viņi salīdzina produkta vai pakalpojuma uztverto vērtību ar cenu. (Almquist, Senior and Bloch, 2016) Ņemiet vērā, ka mārketinga speciālistiem izpratne un analīze par to, kā klienti domā par vienādojuma cenu pusi, ir pazīstams un vienkāršs process. Savukārt stratēģiskajiem vadītājiem saprast, ko patērētāji patiesībā vērtē un kāpēc to novērtē, ir daudz grūtāk. Visbeidzot, uzņēmumiem, kas nodrošina daudzus vērtības komponentus, tiek uzlabota uzņēmējdarbības veikspēja. (Almquist, Senior and Bloch, 2016).

Spēja atpazīt vērtības agrīnā stadijā ir ļoti svarīga uzņēmējdarbības veikspējai. Šis pētījums ir vērsts uz gaisa transporta nozari, kas pēdējā laikā ir piedzīvojusi dziļus satricinājumus pandēmijas un ģeopolitisko nemieru dēļ, kas izraisa jaunus vērtības prasības.

Vērtības būtība pastāvīgi ir atkarīga no skatītāja perspektīvas (Almquist, Cleghorn and Sherer, 2018). Konkursā uzvar tie uzņēmumi, kuri rada visdažādāko, unikālāko, jaunāko un pieprasītāko vērtību piedāvājumu. Tādējādi uzņēmumi saskaras ar izaicinājumu noteikt un paplašināt vērtības piedāvājumu.

Lai veiktu šo uzdevumu, uzņēmumi izmanto klientu aptaujas un konkurences analīzi, lai definētu savu vērtības piedāvājumu. Tomēr vērtību manuāla noteikšana ir laikietilpīgs uzdevums; un iegūto vērtību skaits bieži ir ierobežots.

Lai veiktu šo uzdevumu, uzņēmumi izmanto klientu aptaujas un konkurences analīzi, lai definētu savu vērtības piedāvājumu. Tomēr vērtību manuāla noteikšana ir laikietilpīgs uzdevums; un iegūto vērtību skaits bieži ir ierobežots. Galveno vērtījumu norādošās vērtības tiek formulētas tekstos. Šo datu analīze, izmantojot teksta ieguves metodes, sniedz iespēju vērtību identificēšanai un paplašināšanai (palielinot vērtību skaitu). Automātiskos vērtību atpazīšanas rīkus var uzskatīt par pārlicinošu papildu izaicinājumu uzņēmējdarbības kontekstā.

2. PĒTĪJUMA MĒRĶI UN UZDEVUMI

Lai atrisinātu vērtību identifikācijas automatizācijas problēmu gaisa transporta jomas E-tekstos, ir formulēti šādi pētījuma jautājumi:

1. Kā vērtības jēdziens ir definēts literatūrā?
2. Kāpēc aviācijas nozare prasa vērtību palielināšanu?
3. Vai vērtības var identificēt manuāli?
4. Vai vērtības var kategorizēt?
5. Vai vērtību skaitu var identificēt un paplašināt automātiski?

Lai atbildētu uz šiem izpētes jautājumiem, rodas šādi pētījuma uzdevumi:

Izpētīt literatūru par vērtības jēdziena filozofisko pamatu un faktisko ietekmi

Izpētīt aviācijas nozares vērtības

Noteikt gaisa transporta nozarei specifiskus faktoros, kas ietekmē vērtību attīstību

Veikt eksperimentu, anotējot tekstus, kas raksturo vērtības aviācijas kontekstā

Izstrādāt vērtību kategorizāciju taksonomiju veidā

Atklāt papildu vērtības salīdzinājumā ar tradicionālo vērtību skaitu aviācijas organizācijā

Izveidot automatizētu metodiku vērtību identificēšanai.

3. PĒTĪJUMA MĒRĶIS, OBJEKTS UN PRIEKŠMETS

Pētījuma mērķis ir izstrādāt metodiku automātiskai vērtību identificēšanai un paplašināšanai gaisa transporta jomā.

Pētījuma objekts ir vērtību piedāvājums gaisa transporta jomā.

Pētījuma priekšmets ir Dabiskās valodas apstrādes analīzes metodoloģija, kas ļauj noteikt vērtības kā daļu no vērtības piedāvājuma.

Starpdisciplinārā pieeja pētniecības objektam un tēmai, kuras stratēģijas mērķis ir iegūt jaunas sintētiskas zināšanas par vērtību ieguves veidiem, paver lielas izredzes salīdzinošo zinātnisko pētījumu veikšanai.

4. PĒTNIECĪBAS DARBĪBAS JOMA

Lai izpētītu automātiskas vērtības noteikšanas un paplašināšanas iespēju, šis pētījums sašaurina jomu līdz gaisa transporta nozarei.

Prasība pēc jaunām vērtībām un vērtību skaita palielināšanas (turpmāk šajā darbā vērtību skaita palielināšana tiek dēvēta par vērtību paplašināšanu) Gaisa pārvaldījumos spilgti parādīta Starptautiskās gaisa transporta asociācijas (IATA) nozares lietu komitejas (IAC) ziņojumā "Avioindustrijas nākotne 2035" (IATA, 2018). Dokumenta 12. ieteikumā ir skaidri norādīts, ka, "tā kā jaunās tehnoloģijas un vērtību maiņas maina veidu, kā un kāpēc cilvēki strādā, nozarei būs jāiegulda prasmēs nākamajiem aviācijas vadītājiem un darbiniekiem un jāinformē par ieguvumiem, ko sniedz darbs šajā nozarē".

Arguments par aviācijas plašāku vērtību ir atspoguļots IATA IAC ziņojumā (IATA, 2018) kā Lielo datu, automatizācijas un interneta jautājumu tendenču apvienojums. Paredzams, ka šīs izmaiņas radīs jaunas iespējas, pārveidojot to, kā cilvēki un ierīces tiek monitorētas un mērķētas reāllaikā. Gaisa transporta uzņēmumiem ir izdevīgi gūt labumu no uzlabojumiem automatizācijas, jaunu transporta veidu un patērētāju attieksmes jomā.

Šis pētījums atbalsta koncepciju, ka vērtības datu kontrolei būs priekšrocības salīdzinājumā ar esošajiem konkurentiem jaunu nišu izstrādē. Kā (Almquist, Senior and Bloch, 2016) pierādīts, uzņēmumiem, kas ieguvuši augstus rezultātus četros vai vairāk vērtības elementos, nesena is ieņēmumu pieaugums ir četras reizes lielāks nekā uzņēmumiem ar tikai vienu augstu rezultātu.

Uzņēmumi cenšas attīstīt savu vērtības piedāvājumu, salīdzinot ar konkurentiem un veicot tirgus izpēti. Lai tās paveiktu manuāli, šīs procedūras ir darbietilpīgas un laikietilpīgas. Automatizācijai ir īpašs ieguvums, jo tā palīdz uzņēmumiem noteikt vairāk vērtību, vienlaikus piedāvājot arī jaunas, unikālas un iepriekš neapsvērtas iespējas uzņēmumiem. Ajvari un Žirama (2015) pārbaudīja trīs populārus rīkus vērtību piedāvājumu veidošanai: Ostervaldera un Pignera Value Proposition Canvas (2010), Bārnsa un kolēģu The Value Proposition Builder™ (2009), un Vaildevūra un kolēģu People Value Canvas (2015). Tie visi ir daudzpakāpju rīki, kas ietver tirgus un klientu analīzi.

Tādējādi automatizācijai ir izteikts ieguvums, jo tā palīdz uzņēmumiem atklāt vairāk vērtību, vienlaikus piedāvājot jaunas, unikālas un iepriekš neapsvērtas iespējas uzņēmumiem. Mūsdienā tiešsaistes laikmetā uzņēmumi lielākoties publicē vērtību piedāvājumu, izmantojot savas tīmekļa lapas, tostarp galvenās lapas, tiešsaistes reklāmas un sociālo tīklu profilus. Vērtības piedāvājums tiek sniegts, izmantojot teksta, vizuālās un skaņas tehnikas. Šis pētījums koncentrējas uz tekstuālā vērtību piedāvājumu un izmanto NLP, lai saprastu cilvēku valodu, analizējot tekstu. Teksta ieguve tiek izmantota, lai iegūtu informāciju no nestrukturēta satura.

Vērtību piedāvājuma automatizācija pētījumā ietver tekstus no gaisa transporta jomas jaunuzņēmumu galvenajām lapām un Twitter profiliem. Jaunuzņēmumi tiek izvēlēti kā vērtības datu avots kā daļa no vērtības piedāvājuma saskaņā ar šādiem kritērijiem: jaunuzņēmumi ir vieni no pirmajiem uzņēmumiem, kas norāda jaunas un unikālas vērtības, jo to vērtības piedāvājums ir balstīts uz jaunu biznesa modeli vai inovāciju. Turklāt jaunuzņēmumi savu informāciju par jaunu vērtību sniedz atklāti un apzināti, jo viens no jaunuzņēmumu mērķiem ir nepārtraukta finansējuma piesaiste. Pastāvīgais finansējuma meklēšanas process atspoguļo jaunuzņēmumu būtību: nemitīga izaugsme vērsta globāli. Lai piesaistītu uzmanību un piedāvātu ērtu mijiedarbības veidu, jaunuzņēmumi sniedz informāciju par sevi galveno lapu veidā vienotām datu bāzēm, apkopojot tūkstošiem jaunuzņēmumu.

5. PIEEJA

Vērtības identificēšana (VI) ir kompleksa un sarežģīta problēma. Autore izmanto NLP pieeju, lai apzinātu šai problēmai labi piemērotas metodes un iezīmes. Tiek izmantota uz datiem balstīta tehnika, t.i., autore vispirms izveido datu kopu un pēc tam analizē to, izmantojot skaitļošanas lingvistiskās metodes, lai noteiktu, kuras funkcijas un metodes darbojas vislabāk. Autore izmanto arī korpuslingvistisko pieeju, t.i., uzticama valodas analīze ir vairāk iespējama ar korpusiem, kas savākti laukā tās dabiskajā kontekstā.

6. PĒTĪJUMA METODOLOĢIJA UN METODES

Šajā pētījumā tiek izmantotas kvalitatīvas un kvantitatīvās metodes, lai izveidotu metodiku automātiskai vērtību identificēšanai gaisa transporta jomā. Pirmais solis ir vērtību teorētiskā analīze, kuras mērķis ir noteikt vērtību identificēšanas un paplašināšanas procesa metaipašības gaisa transporta jomā. Tā ietver sekundāro datu analīzi: teorētiskās literatūras, iepriekšējo pētījumu, politikas dokumentu un teorētiskās struktūras pārskatīšanu kā pieeju un jēdzienu novērtēšanu un konceptualizāciju. Otrais solis bija pētniecības procedūras izstrāde, tāpēc tika pielietotas mūsdienīgas pieejas informācijas ieguvei un sistematizācijai: semantiska anotācija apvienojumā ar binārās klasifikācijas pieeju, beztermiņa apsekojumu, klasifikāciju un organizēšanu taksonomijās. Rezultātu īstenošana tika pārbaudīta, izmantojot klasifikatoru: neironu tīklu.

7. PĒTĪJUMĀ ZINĀTNISKĀ NOVITĀTE

Pētījuma novitāte slēpjas vismodernākajā metodoloģijā, kas ļauj teksta ieguves metodes kā daļu no dabiskās valodas apstrādes (NLP) piemērot daudzajām jaunuzņēmumu galvenajām lapām, lai apkopotu datus kā daļu no vērtības noteikšanas un paplašināšanas automatizācijas procesa, detalizējot:

1. Tika piedāvāts jauns uzdevums - vērtību identificēšana no tekstiem, un tiek ieviesta jauna Nosauktā vienība: Vērtības vienības atpazīšana (VER).
2. Izstrādāta metodika vērtību identifikācijas automatizācijai no tekstiem ar mērķi palielināt vērtību skaitu.
3. Radīts vērtību korpus/datu kopa gaisa transportam.
4. Vērtību sistematizācija: taksonomiju veidā tika izstrādātas divas vērtību kategorizācijas:
 - Gaisa transporta vērtības taksonomija, klientu perspektīva.
 - Vērtību semantiskā taksonomija aviācijā, nozares perspektīva.
5. Tika izstrādāts konceptuāls ietvars:
 - Vērtības izteiksme: izstrādāta viena vārda koncepcija.
 - Vērtības sociālie aspekti: izstrādāta individuālā un grupas vērtību attiecības koncepcija.
 - Tika noteikta gaisa transporta vērtību maiņa vēsturē.
 - Veikta termina vērtības piedāvājums etimoloģiskā ģenēze.
 - Noteikti galvenie faktori, kas ietekmē vērtības piedāvājumu aviācijā.

8. PĒTĪJUMA TEORĒTISKĀ UN PRAKTISKĀ VĒRTĪBA

Šis pētījums ir pirmais pētījums, kurā vērtību kvantitatīvās īpašības kā vērtību piedāvājuma daļa tiek aplūkota kā primārā izpētes joma. Turklāt šajā pētījumā izmantota jauna pieeja, lai jaunuzņēmumu tekstus izmantotu kā pētniecības bāzi vērtību noteikšanai un paplašināšanai. Tas paplašina esošās metodes uzņēmuma vērtības piedāvājuma izstrādei. Turklāt iegūtie dati liecina, ka vērtības var būtiski paplašināt un klasificēt, lai atbalstītu lēmumu pieņemšanas procesu. Tādējādi tika konstatēts, ka dziļās mācīšanās rīkus var izmantot, lai automatizētu vērtību identificēšanas un paplašināšanas procesu.

Pētījuma teorētisko ieguldījumu var formulēt šādi:

1. Šis pētījums sniedz jaunu koncepciju par iespēju automatizētai vērtību identificēšanai un paplašināšanai, kas noved pie iespējamās kategorizācijas.
2. Šis pētījums ir pirmais, kas veic automatizētas vērtības noteikšanas un paplašināšanas uzdevumu. Tas nodrošina teorētisko sistēmu automatizētai vērtības identifikācijai un paplašināšanai, ko var piemērot jebkurā jomā.
3. Pētījumā ir izskaidrots, kādi faktori jāņem vērā vērtību automatizētā identificēšanā un paplašināšanā kā daļa no vērtības piedāvājuma: vērtības jēdziens, anotatoru savstarpējā vienošanās, viena vārda pieeja klasifikācijā, klasifikatora precizitāte.
4. Visbeidzot, sniegtās vērtību kategorijas taksonomiju veidā veicināja sekojošo: kopīga valoda konceptu koplietošanai, zināšanu atklāšanai (modeļi, tendences), ļaujot atkārtoti izmantot informāciju, nevis rekreāciju, saskaņot priekšmetus vai saturu, kas koplieto tos pašus jēdzienus, nodrošināt metadatus identifikācijai, salīdzināšanai, analīzei.

Turpmākajos paziņojumos ir uzsvērtā pētniecības praktiskā nozīme:

1. Stratēģisko lēmumu pieņemšanas darbības joma tiek paplašināta, nodrošinot paplašinātu vērtību izvēli, pamatojoties uz inovācijām un jauniem uzņēmējdarbības modeļiem..
2. Taksonomijas ļauj izstrādāt kopīgu valodu jaunām koncepcijām; nodrošina saskanīgus datus identifikācijai, salīdzināšanai un analīzei.
3. Automatizācija sniedz iespējas identificēt nozares vērtību maiņu laikā, un jaunas vērtības var dokumentēt, izmantojot laika un resursu taupīšanas metodi.
4. Jaunajai metodoloģijai ir potenciāls dažādās jomās, kas ļauj identificēt un paplašināt vērtības ārpus gaisa transporta.

9. PĒTĪJUMA APPROBĒŠANA

Pamatjēdzieni un konstatējumi tika paziņoti un apspriesti zinātniskās konferencēs un semināros. Autorei ir 12 publikācijas par zinātniskā darba tēmu un dalība 23 konferencēs, darbnīcās un reālās dzīves biznesa procesos Latvijā, Japānā, Itālijā, Igaunijā un Lietuvā. Šajā pētījumā gūtie sasniegumi un konstatējumi tika

izmantoti Transporta un telekomunikāciju institūta pedagoģiskajā praksē kā gadījumu izpēte, kas iekļauta tēmā “Digitālais mārketing”.

10. PROMOCIJAS DARBA STRUKTŪRA

Darbu veido 5 nodaļas un 14 pielikumi. Tajā ir 194 lappuses, 39 attēli un 24 tabulas. Atsauču un informācijas avotu sarakstā ir 301 nosaukumu.

Ievadā ir izklāstīta pētījuma motivācija, mērķis, uzdevumi, objekts un subjekts, pētījuma novitāte, empīriskais un teorētiskais ieguldījums, kā arī tā struktūra.

Pirmajā nodaļā ir aprakstīta aviācijas stratēģiskās vadības attīstība, jo gaisa transports šajā pētījumā tiek ņemts par lietojuma gadījumu un Vērtības piedāvājuma jēdziens un vērtības kā daļa no šīs koncepcijas. Vērtības un vērtības piedāvājums šajā kontekstā ir šī pētījuma objekts. Šī pētījuma datu avots ir jaunuzņēmumu tiešsaistes teksti. Šajā nodaļā aplūkots ieguvums, izmantojot jaunuzņēmumu tiešsaistes tekstus par datu avotu.

Otrajā nodaļā aplūkotas uz datiem balstītas lēmumu pieņemšanas un lēmumu atbalsta sistēmas, uz datiem balstītu lēmumu pieņemšanas priekšrocības un sniegts pētījuma metodoloģiskais pamatojums. Tiek prezentētas informācijas ieguves metožu iespējas biznesa procesu atbalstam. Informācijas ieguves metožu sašaurināšanās līdz nosaukto vienību (entity) atpazīšanas un binārās klasifikācijas pieejām, jo ir parādītas dažas no galvenajām šī pētījuma teksta ieguves metodēm. Vērtību vienības (value entity) atpazīšana tiek piedāvāta kā rezultātā kombinēta metode.

Trešajā nodaļā aprakstīta eksperimenta norise, taksonomijas izstrādes principi un jauna viena vārda pieeja vērtību identificēšanai ar anotatoriem un klasifikatoriem.

Ceturtajā nodaļā sniegtas galvenās atziņas un atbildes uz pētījuma jautājumiem: konceptuālā analīze, vērtību korpuss, vērtību hierarhiskās klasifikācijas, iespēja automatizēt vērtību identificēšanu un paplašināšanas procesu, izstrādājot dabiskās valodas apstrādes modeli.

Piektajā nodaļā ir sniegts zinātniskā darba empīriskais un teorētiskais ieguldījums un diskusija par vērtību analīzi un rezultātu praktisko nozīmi.

Secinājumos ir uzskaitīti galvenie pētījuma rezultāti un izceltas svarīgākās atziņas. Tiek parādīts, ka aizstāvībai izvirzītie apgalvojumi ir apstiprinājušies. Visbeidzot, tiek izcelta perspektīva turpmākiem pētījumiem un sniegti ieteikumi nozares analītiķiem un lēmumu pieņēmējiem.

11. AIZSTĀVĒŠANAI IZVIRZĪTĀS TĒZES

Aizstāvībai izvirzīti šādi apgalvojumi:

1. Vērtības var noteikt tekstos (galvenās lapas, Twitter profili)
2. Vērtības gaisa transporta jomā var iedalīt kategorijās, izmantojot taksonomijas

3. Metodoloģija ļauj automātiski identificēt vērtības no tekstuālas informācijas par uzņēmumu

12. PROMOCIJAS DARBA NODAĻU KOPSAVILKUMS

12.1. Vērtības kā vērtības piedāvājuma daļa gaisa pārvadājumos

Vērtības kā daļa no vērtības piedāvājuma koncepcijas

Uzņēmumi pieņem jaunas vērtības, lai palielinātu rentabilitāti un noturību. Saskaņā ar (Almquist, Senior and Bloch, 2016) šiem vērtības pētījuma elementiem, uzņēmumu panākumus veicina spēja sasniegt vairākus Vērtības elementus. Savos pētījumos Almkvists un citi sniedz piemērus par pasaules vadošo viedtālrunu zīmolu vērtību skaitu: piemēram, LG piegādā 3 elementus, Samsung 10 un Apple 12; uzņēmums, kas nodrošina labu sniegumu lielākam skaitam elementu nekā citi tirgus dalībnieki, var piesaistīt vairāk klientu un var pieprasīt augstāku cenu nekā konkurenti. Piemēri: Red Bull - uzņēmums sponsorē ekstrēmās sporta veidus, kas savieno sporta entuziastus visā pasaulē, Nordstrom (uzņēmums iegādājās Trunk Club, radot saistības un piederības sajūtu).

Pirmo (Lanning un Michaels, 1988) pirmo reizi izmantojot izteicienu "Vērtības piedāvājums" (VP) 1988. gada darba dokumentā konsultāciju uzņēmumam "McKinsey and Co", (Almquist, Senior and Bloch, 2016) ierosināja oriģinālu stratēģiju, kuras pamatā ir diferencēts klienta vērtības piedāvājums. Ieteikto vērtību kopu sauca par Vērtības elementiem. Vērtības elementi bija kategorijas, kuru pamatā ir klasiskā Maslova vajadzību hierarhija (Maslow, 1943).

Maslova vajadzību hierarhija (Maslow, 1943) līdz šim ir kalpojusi par pamatu cilvēka būtības un cilvēka labklājības prasību novērtēšanai. Pirmo reizi (Almkvists, seniors un Blohs, 2016) tika pieņemta šo cilvēcisko vajadzību hierarhija to vērtību hierarhijā, kuras klienti sagaida no biznesa. Šī ideja kalpo kā spilgts piemērs biznesa un zinātnes sadarbībai.

Ir salīdzinoši maz pētījumu par vērtību priekšlikumu analīzi tiešsaistes jaunuzņēmumos. (Su-C Li, 2007) argumentē, ka pareizi izstrādāts vērtības piedāvājums ir būtisks vērtības radīšanas procesā e-biznesā un vērtību kopražošana ir vērtību aizsardzības mehānisma pamatelementi tīkla ekonomikā.

Ajvari un Žirama (Äyväri and Jyrämä, 2017) savā rakstā "Vērtību piedāvājuma rīku pārdomāšana dzīvajām laboratorijām" sniedz konceptuālu analīzi par vērtību piedāvājuma rīkiem, ko izmantot turpmākajos empīriskajos pētījumos un veidojot pārvaldības ieskatu. Konceptuālās analīzes kontekstā tiek apskatītas trīs vadības metodes vērtību piedāvājumu izstrādei, kas balstās uz dzīvo laboratoriju ietvariem un pašreizējiem teorētiskajiem sasniegumiem saistībā ar vērtības jēdzienu. Starp atklājumiem dzīvo laboratoriju pieejas kontekstā, the Value Proposition Builder šķiet pretrunā ar uz lietotāju orientētu inovācijas procesu idejām un telpām (Äyväri and Jyrämä, 2017).

Šis pētījums vērtības kontekstā pieņem šādus terminus kā "paziņojums, kurā skaidri norādīts, kādu labumu klients saņems, iegādājoties konkrētu produktu vai

pakalpojumu” (Rintamäki et al, 2007): Vērtība, Vērtības piedāvājums, Klienta vērtība, Klienta vērtības piedāvājums.

Konkurences priekšrocības pēc nesenās liberalizācijas rada jaunas problēmas; aviācijas nozarei būtu jāņem vērā prasība identificēt un nodrošināt pasažieru vērtības, kā arī nodrošināt iekšējo vērtību. Aviācijā tradicionālās dominējošās vērtības ir cena, komforts un ātrums, un tās dominē jau gadu desmitiem (IATA Passenger Surveys, 1991-2019; Bain&Company, 2011).

1. tabulā apkopoti faktori, kas ietekmē vērtības noteikšanu aviācijā. Šajā tabulā norādīts faktors, tā ekonomiskās īpašības un īpašā ietekme uz vērtību.

1. tabula. Faktori, kas ietekmē vērtības izvirzīšanu aviācijā (Autore)

Faktors	Ekonomiskie rādītāji	Ietekme uz vērtību
Produkts tūlītējai lietošanai	aviācijas produkts pēc būtības ir nemateriāls, tūlītējai lietošanai, netiek uzglabāts	īstermiņā nav segti jaudas nodrošināšanas izdevumi; spiežot uz cenu samazināšanu
Līdzīgi produkti	aviokompānijas produkts pēc būtības ir viendabīgs, līdzīgs visās aviosabiedrībās	jauno produktu iniciatīvu ātri atdarina konkurenti; nav diferenciacijas vērtības
Zema robežizmaksu struktūra	zemas robežizmaksas, augstas fiksētās izmaksas papildu pasažieriem	atkarība no augstām fiksētām izmaksām ierobežo vērtības piedāvājuma diapazonu
Būtiskas izejas barjeras	ilgtermiņa gaisa kuģu un lidostu kapacitāte; stingrs pilna servisa tīkla operatoru biznesa modelis	fiksēts negrozāms vērtības piedāvājums; jebkuri papildu veidi esošo resursu izmantošanai ir izdevīgi (piemēram, tikties pēc jaunām vērtībām)
Jauda	Jaunu lidmašīnu un lidostu jauda; ilgtermiņa ieguldījumu atdeve	jauda ir fiksēta / neelasīga problēma: nepietiekama izmantošana (ja nepietiekams pieprasījums) nespēja viegli palielināt jaudu
Nozares izaugsme	nepastāvīga un neviendabīga	nestabili periodi izraisa konservatīvu pieeju vērtību piedāvājumam
Nosacījumu neviendabīgums	lidostas un gaisa pārvadātāji ir pakļauti īpašai politikai to vietējā tirgū	tas ietekmē konkurētspējīgu sadarbību starp uzņēmumiem veidā, kas nav saistīts ar vērtības piedāvājumu vai efektivitāti; tāpēc nav nepieciešams meklēt jaunas vērtības
Nesenā atbrīvošanās, monopolizācija	mazāka konkurence	Vērtība kā pēdēja

Jaunizveidots uzņēmums: koncentrēta un atvērta vērtību piedāvājuma avots

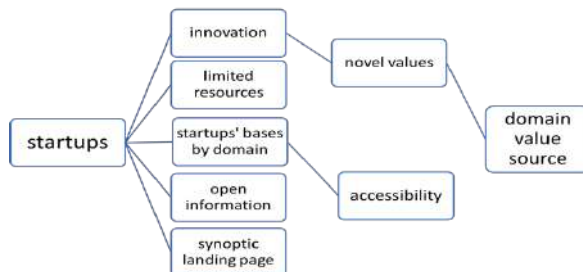
Eiropas Komisija jaunuzņēmumus definē kā “bieži ar tehnoloģijām balstītus uzņēmumus, kas kopumā apvieno ātru izaugsmi, lielu paļaušanos uz produktu, procesu un finansējuma jauninājumiem, īpašu uzmanību pievēršot jaunu tehnoloģiju attīstībai un plašu inovatīvu uzņēmējdarbības modeļu izmantošanu, un bieži vien arī sadarbības platformas” (European Commission, 2016).

Mūsdienās arvien vairāk uzņēmējdarbības vidi, uz datiem balstītu inovāciju un ekonomiskās vērtības radīšanu rada vairākas organizācijas, nevis viena organizācija vai parasta vērtību ķēde (Hein et al, 2019). Kā aizstājēju dažādu datu kombinācijas veido daudznozaru datu bāzes, bagātinot uzņēmumu starpnozaru un

sociālo tehnoloģiju potenciālu (Hein et al, 2019; Oliveira and Loscio, 2018; Yoo et al, 2010). Vairākus jaunuzņēmumus no vienas nozares jomas var uzskatīt par ekosistēmu, kas savstarpēji sniedz informāciju par vērtības noteikšanu.

Tā kā jaunuzņēmumu mērķis ir strauji augt un mērogoties (Giardino et al, 2014) ar “lielu paļaušanos uz produktu, procesu inovācijām” (European Commission, 2016), informācija par to galvenajām lapām ir atklāta un faktiska; tas nodrošina labu datu avotu, lai analizētu esošo un perspektīvo vērtības piedāvājumu.

1. attēlā apkopotas jaunuzņēmumu iespējas kļūt par vērtību avotu noteiktai jomai.



1. attēls. Jaunuzņēmumi kā vērtību avots (autores veidots)

Nobeiguma piezīmes

1. Aviācijas attīstības korporatīvajā līmenī lēmumi tiek pieņemti, izmantojot tādu instrumentu kā stratēģiskā vadība.
2. Daudzi stratēģijas jomas eksperti iekļauj vērtības radīšanu kā galveno veiksmīga uzņēmuma stratēģijas potenciālo sastāvdaļu.
3. Ņemot vērā pieaugošo konkurenci un gaisa transporta liberalizāciju, vērtības piedāvājums aviācijā pieaug un strauji kļūst par stratēģisku prioritāti.
4. Tradicionālās vērtības gaisa pārvadājumos ir cena, kvalitāte un ātrums. Pēdējās desmitgadēs gaisa transporta komercializācija aizsāk konkurenci un līdz ar to arī jaunu vērtību meklējumus.
5. Vērtības radīšana ir viens no galvenajiem uzņēmuma mērķiem, un pēc nesenās liberalizācijas gaisa transporta nozare saskaras ar neizpildītu vajadzību noteikt jaunu vērtības piedāvājumu. Šī procesa automatizācija būtu izdevīga.

Lai automatizētu vērtības atpazīšanas procesu, jaunuzņēmumi uzrāda pienācīgu datu avotu vērtības piedāvājuma analīzei, jo to mērķis ir strauji augt un mērogoties, informācija galvenajās lapās ir atklāta un faktiska.

12.2. Metodoloģiskās pieejas: Pašreizējais stāvoklis/State of art

Tradicionālās, uz datiem nebalstītās pieejas lēmumu pieņemšanā, radot vērtības piedāvājumu, piemēram, Ostervaldera un Pigneura piedāvātās (2010. gads), parasti izmanto no datiem neatkarīgu, manuālu pieeju, iesaistot vienu vai dažus dalībniekus.

Lai modelētu, kā uzņēmumi īsteno savus uzņēmumus, tika izveidotas dažādas biznesa modeļu (BM) sistēmas, piemēram, plaši atzītā Business Model Canvas,

tostarp vērtības piedāvājuma izveide, kas atbalsta vispārpieņemtu biznesa modeļa interpretāciju uzņēmumā. Tomēr šiem tipiskajiem, uz datiem nebalstītajiem lēmumu pieņemšanas procesiem ir vairāki trūkumi, kas ierobežo to lietderību.

Informācijas ieguve

Līdz ar globālā tīmekļa eksplozīvo izaugsmi, atzīst (Chen, 2014), ka tīmeklī ir pieejams plašs informācijas klāsts daļēji strukturētos un nestrukturētos dokumentos (piemēram, tīmekļa lapās un pilna teksta dokumentos). Tādējādi pēdējos desmit gadus turpinās centieni automātiski piekļūt informācijai no šiem avotiem un to novērst. Šo apgabalu sauc par Informācijas ieguvu (IE). Atšķirībā no Informācijas izguves (IR), kuras uzdevums ir noskaidrot un klasificēt attiecīgos dokumentus atbilstoši vaicājumiem, kas atspoguļo lietotāja informācijas nepieciešamību, IE sistēmu plašs mērķis ir iegūt strukturētus datus no daļēji strukturētiem un/vai nestrukturētiem mašīnlasāmiem dokumentiem, ja iegūstamās informācijas konkrētais veids un struktūra ir atkarīga no konkrētā lietojuma nepieciešamības. Web IE rīki, saukti arī par apvalka indukcijas sistēmām/wrapper induction systems, (piem., Lixto (Baumgartner, 2001), RoadRunner (Crescenzi et al, 2001) un WebTable (WebTable, 2022)) veic IE ar daļēji strukturētiem tiešsaistes dokumentiem, piemēram, HTML lapām, galvenokārt izmantojot dokumentu izkārtojuma struktūru. Turpretī IE uzdevumi ir tie, kas apstrādā brīvo tekstu dabiskā valodā, jo šie “nestrukturētie” dati ir primārais cilvēka radītas informācijas avots (Sarawagi, 2006).

Saisītais darbs ar nosaukto vienību/entity atpazīšanu

Šajā pētījumā ierosināts jauns vērtības noteikšanas uzdevums vērtības noteikšanas ietvaros kā daļa no esošajām Nosauktās vienības atzīšanas/Named Entity Recognition metodēm (NER), kur vārdu vai frāzi, kas izsaka vērtību, uzskata par Nosaukto vienību/Named Entity.

Šajā nodaļā aprakstītā informācijas ieguves paņēmiena galvenā joma ir NER, un svarīgākie saistītie darbi ir apkopoti šajā sadaļā.

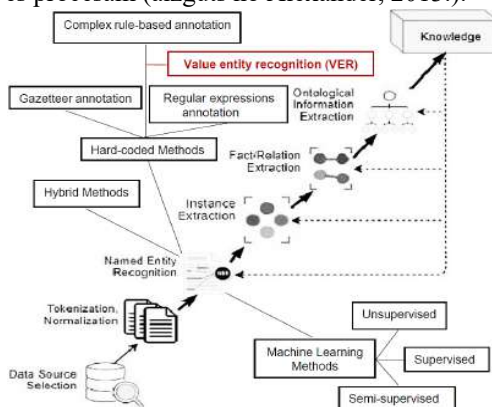
(Ortona, 2016) secina, ka pašlaik varētu noteikt trīs pētījumu grupas Nosaukto vienību atpazīšanai, pamatojoties uz informācijas avotu, kas tiek izmantots teksta anotācijas sagatavošanai: *uz valodu orientēta pieeja* atspoguļo tradīcijas nosaukto vienību atpazīšanā. Tā izmanto dabisko valodu zināšanas, lai iegūtu valodu modeli, kas parasti veido noteikumus, kurus, piemērojot tekstiem, izgūst interešu vienību instances; *uz zināšanām balstīta pieeja* izmanto strukturētu, t. i., ontoloģisku vai nestrukturētu, t. i., ievāktu pamatinformāciju, kas atrod interesējošo vienību teksta dokumentos; *anotētāju apvienošana*. Parasti pieejas, kas attiecas uz vairāku NER sistēmu apkopošanu, ir izmantojušas balsošanas mehānismu vai tā variācijas (Kozareva, 2007).

Parasti šie paņēmieni bieži vien nepievērš uzmanību noteiktai informācijai par teksta semantiku vai kontekstu. Piemēram, proti, Nerds (Nerd, 2022), pierādīja, ka ontoloģiskās zināšanas par fonu - tas varēja noderēt, veidojot anotācijas - tikpat labi varētu noderēt, apkopojot vairākus anotētājus, loģiski precizējot nesakarīgas anotācijas.

Vērtības vienības atpazīšana/Value Entity Recognition (VER)

Parasti nosauktās vienības ir īpašvārdu veids, bet citas ir vispārīgas vienības domēnā. Tādējādi NER tiek apstrādāts, pamatojoties uz ortogrāfiskām pazīmēm, vārda līmeņa pazīmēm un ģeogrāfisko nosaukumu rādītājiem (Ngo, 2021). Lai klasificētu nosaukto vienību tagus un nosaukto vienību atdalīšanu, pamatojoties uz to iezīmēm, tiek izmantotas datu analīzes metodes. Tomēr šīm pazīmēm ir ļoti mazs ieguldījums. Vērtības vienības atpazīšanas (VER) uzdevumā, jo, kā parādīja (Ngo, 2021), pamatojoties uz lauksaimniecības vienību atpazīšanu (AGER), vispārīgās vienības nav īpašvārdi un tām nav ortogrāfisku pazīmju, piemēram, nosauktās vienības NER uzdevumā. Vienību atpazīšanas sistēmām tas ir izaicinājums.

Šī disertācija piedāvā vienības atpazīšanas procesu noteiktā sfērā, kas ir gaisa transports, un vienības ir vērtība kā vērtības piedāvājuma daļa. Būtiskākie VER fakti ir: (1) noformējuma likumos balstīts anotācijas eksperiments vērtību vienībām; (2) izveidot semantiski balstītu Vērtību vienības atpazīšanas (VER) uzdevumu un oriģinālu efektīvu pieeju VER gaisa transporta jomā; (3) izmantot padziļināto mācīšanos (DNN), lai apmācītu un identificētu vērtību vienības no jaunuzņēmumu gaisa pārvadājumu Twitter profilu tekstiem. Semantiski balstītā VER pieeja apvieno uz NER balstītu pieeju un binārās klasifikācijas pieeju (ja vienība ir vērtība vai nav vērtība). 2. attēlā parādīta vērtību vienības atpazīšanas (VER) vieta kā NER pieeja informācijas ieguves procesam (aizgūts no Alexander, 2013.).



2. attēls. Vērtību vienību atpazīšana (VER) kā NER pieeja informācijas ieguves procesam (pamatojoties uz Alexander, 2013)

Nosakot metodes vērtību noteikšanai no tekstiem, tika izvērtētas vairākas alternatīvas, kas atspoguļotas 2. tabulā.

2. tabula. Alternatīvās metodes vērtību identificēšanai no tekstiem

Informācijas iegūšanas metode	Pienemšanas / noraidīšanas iemesls
1. Daudzklašu (ātrums, izmaksas, komforts), iepriekš noteiktas klases	1. daudzas klases un nav iespējas atrast jaunas klases/vērtības
2. sentimenta/noskaņojuma analīze	2. tiek vērtēts viss teksts/teikums; Daudzklašu (vērtības var piederēt sentimentu klašu sajaukumam)
3. binārā terminu klasifikācija	3. atļauj identificēt jaunas vērtības
4. Nosauktās vienības atpazīšana (NER)	4. atļauj izveidot jaunu vienību atpazīšanai

NER pieeja ļauj izveidot *jaunu vienības* tipu - *vērtību* ar iepriekš definētiem semantiskiem noteikumiem anotatoriem: viņi apgūst vērtības piedāvājuma pamatteoriju un ir apmācīti pavadīt tikpat daudz laika, cik vidusmēra tiešsaistes patērētājs pavada tīmekļa lapā - 62 sekundes (Contentsquare Homepage, 2020).

Noslēguma piezīmes

1. Tradicionālajai taktikai, kas nav balstīta uz datiem, ir vairāki trūkumi, kas ierobežo lēmumu pieņemšanas efektivitāti: cilvēkresursu ieguldījums, aizspriedumi un subjektivitāte, kas ir pakļauti cilvēku lēmēju kļūdām, parasti ir dārgāka, laukietilpīga..
2. Lai atbalstītu biznesa procesu jauna vērtības piedāvājuma sniegšanai, šis pētījums piedāvā uz datiem balstītu pieeju vērtību identificēšanai, vērtību skaita palielināšanai un šī procesa automatizācijai.
3. Uzņēmuma vērtības piedāvājums vairumā gadījumu ir izteikts dabiskajā valodā un satur vērtības, kas izteiktas ar vārdiem.
4. Ar aviāciju saistītu darbu analīze angļu valodā norāda uz pētniecības trūkumu šādās jomās: teksti no tīmekļa lapām, oficiāliem noteikumiem, uzņēmumu politikām, darba ziņojumiem, izpētes dokumentiem Gaisa transporta jomā, kas izmantojami kā datu bāzes vērtību identificēšanai un numuru palielināšanai.
5. No Dabiskās valodas apstrādes viedokļa, vērtības vērtību piedāvājumu tekstos var uzskatīt par nosauktām vienībām un identificēt, izmantojot metodi, ko sauc par Nosaukto vienību atpazīšanu.

Šī disertācija piedāvā vienības atpazīšanas procesu noteiktā sfērā, kas ir gaisa transports, un vienība ir vērtība kā vērtības piedāvājuma daļa. Būtiskākie VER fakti ir: (1) noformējuma likumos balstīts anotācijas eksperiments vērtību vienībām; (2) izstrādāt semantiski balstītus vērtību vienību atpazīšanas (VER) uzdevumus un jaunu efektīvu pieeju VER gaisa transporta jomā; (3) izmantot padziļināto mācīšanos (DNN), lai apmācītu un identificētu vērtību vienības no gaisa transporta jaunuzņēmumu Twitter profilu tekstiem. Uz semantisku balstītā VER pieeja apvieno uz NER balstītu pieeju un binārās klasifikācijas pieeju.

12.3. Metodoloģija un pētījumu dizains

Šajā nodaļā ir apskatīta metodoloģija un pētījuma plāns. Prezentētais pētījums ir jauns no esošo metožu kopuma viedokļa, kas pielietots jaunam uzdevumam: identificēt un paplašināt vērtības gaisa pārvadājumos un parādīt iespējas automatizēt noteikšanas un paplašināšanas procesu. Jaunums “Viena vārda princips”/“*One-Word principle*” ir izmantots datu analīzei.

Pētījuma pieeja un metodes

Pētījuma procedūra ir sadalīta *septiņos soļos*:

1) posms ir vērtību teorētiskā analīze, kuras mērķis ir noteikt vērtības noteikšanas un paplašināšanas procesa metaīpašības gaisa transporta jomā. Tas ietver: sekundāro datu analīzi kā teorētiskās literatūras, iepriekšējo pētījumu, politikas dokumentu pārskatīšanu; koncepciju veidošanu kā pieeju un jēdzienu izvērtēšanu un konceptualizāciju;

2) solis nosaka gala nosacījumus, kas jāievēro: tiek atpazītas un kategorizētas vārdos izteiktas vērtības, panākta ekspertu vienošanās un izveidots modelis;

3) posmā tiek izmantotas divas pētniecības metodes: **empīriskā līdz konceptuālai un konceptuālā līdz empīriskai.**

Empīriskā līdz konceptuālai

4a) lai identificētu objektu apakškopu, pētniecības bāzi nosaka kā jaunuzņēmumu galvenās lapas gaisa transporta jomā; vērtības tiek anotētas, datu kopa tiek veidota kā uz izlasi vērsta pieeja bez iepriekšējas konceptualizācijas.

5a) lai identificētu kopīgas īpašības un grupētu objektus, tiek veikta vērtību semantiskā analīze, kuras mērķis ir potenciāla kategorizācija.

6a) lai grupētu īpašības dimensijās taksonomijas izveidei, kategorijas tiek pārbaudītas un hierarhiski sakārtotas apakšklasēs un pēc tam klasēs.

Konceptuālā līdz empīriskai

4b) objektu īpašību un izmēru iepriekšēja konceptualizācija: tiek pieņemts, ka vērtības tekstos var identificēt un paplašināt automātiski.

5b) lai pārbaudītu objektus pēc šīm īpašībām un izmēriem, tika profesionāli sagatavots neironu tīkla modelis.

6b) tiek saņemta galīgā datu kopa un aprēķināta precizitāte.

7) apstiprināt, ka ir izpildīti galīgie nosacījumi.

Lai izstrādātu pētniecības procedūru, tika pārbaudītas mūsdienīgas pieejas informācijas iegūšanai no tekstiem un izdarīts šāds secinājums: metadatus, ko iegūst IE uzdevumi, lai saprastu neapstrādātu tekstu, uzskata par semantisko anotāciju, un teksta analīzes izpildītāji - cilvēki vai programmas - risina IE uzdevumus kā semantiskie anotētāji.

Kvalificējot vērtības vienību atpazīšanu kā semantisku anotāciju veicēju uzdevumu, bija jāanalizē dažādu IE pieeju potenciāls jaunam vērtības noteikšanas uzdevumam:

- *Nosaukto vienību ieguve*: atpazīt un klasificēt atomiskus fragmentus tekstā kā tādus, kas attiecas uz noteiktām semantiskām kategorijām (piem., Person and Company) (Cunningham, 2002). Dažos gadījumos tas ietver arī apakšuzdevumu ar nosaukto vienību noskaidrošana, kas atrisina konstatētās vienības instances vienību identifikatorus (Li et al, 2013);
- *Saistību noskaidrošana*: noskaidrot saistības starp vienībām (Reichartz et al, 2010);
- *Sentimenta/noskaņojuma anotācija*: noteikt viedokļus vai emocijas pamatā esošajā tekstā. Attieksme vai tonis var būt polaritātes viedokļi, piemēram, pozitīvs/negatīvs, vai salīdzinoši viedokļi, piemēram, labāks/sliktāks vai vienkārši subjektīvs/objektīvs (Ganapathibhotla and Liu, 2008);
- *Tēmas anotācija*: pievienot dokumentus ar tēmām vai tematiem (Silva et al, 2001);

- *Līdzsauceles anotācija*: uztvert konkrētas vienības līdzsauceles pieminējumus (Deemter and Kibble, 1995);
- *Atslēgas vārdu anotācija*: izgūt teksta saturam atbilstošus atslēgvārdus vai terminoloģiju (Liu et al., 2009).

No Dabiskās valodas apstrādes viedokļa vērtības vērtību piedāvājumu tekstos potenciāli var uzskatīt par nosauktām vienībām un identificēt, izmantojot metodi, ko sauc par nosaukto vienību pa. Pēc tam, kad tika rūpīgi izvērtēts, vai vērtības ir vērtības piedāvājuma daļas (sk. 2.5. punktu), tika ieviests jauns vienību ieviešanas veids – vērtības vienību atpazīšana (VER).

Iepriekš aprakstītās *pieejas* ir ļāvušas izvēlēties **datu vākšanai izmantojamās metodes**:

- Semantiskā anotācija apvienojumā ar binārās klasifikācijas pieeju (ja vienība ir vērtība vai nav vērtība) galveno lapu marķēšanai; Twitter profili jaunuzņēmumu marķēšanai.
- Aptauja – izriet no pašu formulētajām atbildēm uz atklāto jautājumu *Kādas vērtības jūs sagaidāt no gaisa transporta nozares?*

Iegūto datu sistematizēšanai tika izmantotas šādas **datu analīzes metodes**: kategorizēšana un organizēšana taksonomijās.

Vārdu korpuss veido nekontrolētu vārdu krājumu, piemēram, nav klasificēts un nav sakārtots datu izguvei. Lai vārdu krājumu izmantotu tēzauros, ontoloģijās, taksonomijās un citās zināšanu organizācijas shēmās, sākotnējais nekontrolētais vārdu krājums tiek sakārtots *kontrolētā vārdu krājumā* - kontrolēto vārdu krājumu parasti izmanto satura aprakstīšanai, piešķirot terminus, lai attēlotu metadatus, kas saistīti ar satura objektiem, organizējot saturu tīmekļa vietnēs un līdzīgi (University of Texas, 2022).

Ir četri svarīgi vārdu krājuma kontroles principi, kas virza to uzbūvi un attīstību. Lai izveidotu vadāmus vārdu krājumus, pastāv četri galvenie principi, kas palīdz izstrādāt un attīstīt procesu: *novērstu neskaidrības, kontrolētu sinonīmus, izveidotu terminu sakarības, testētu un validētu* (ANSI/NISO Z39.19-2005). Kontrolētos vārdu krājumos ir iespējams identificēt daudzas semantiskās saiknes starp terminiem. Tie ir saistīti ar asociatīvām attiecībām, hierarhiskām attiecībām un vienlīdzības savienojumiem, kurus var aprakstīt atbilstoši konkrētās īstenošanas vajadzībām.

Eksperimenta procedūra

Lai identificētu un paplašinātu vērtības no gaisa transporta jaunuzņēmumu tekstiem, tika veikts eksperiments, kas sastāvēja no **2 posmiem**. 1. posms ietver aptauju un jaunuzņēmumu anotācijas (3. attēls), bet 2. posms ietver 1. posma jaunuzņēmumu Twitter profilu anotāciju. Galvenā izmantotā tehnika ir anotācija. Datu anotācija ir teksta ieguves tehnika, ar kuras palīdzību dati tiek marķēti, lai objektus (šajā gadījumā - vērtības, kas izteiktas ar vārdiem) padarītu atpazīstamus mašīnām.

Pētījumu bāze

Eksperiments tiek veikts par jaunuzņēmumu tekstiem kā vērtību datu avotu.

Eksperiments, 1. posms. 796 jaunuzņēmumu galvenās lapas gaisa transporta jomā, avots: www.crunchbase.com.

Eksperiments, 2. posms. 796 jaunuzņēmumu Twitter profili no gaisa transporta jomas 1. posma, avots: www.crunchbase.com

1.posms: Izglītība, priekšizpēte, apmācība, anotācija, pēcaptauja (3. attēls).



3. attēls. Procedūras pirmais posms

Anotētāji tika izglītoti par vērtības noteikšanas koncepciju, veica aptauju ar atvērtu vērtību formulējumu AT, apmācīti marķēt vērtības, marķēt vērtības un veikt pēcaptauju, lai atrastu atšķirību no priekšizpētes turpmākajiem pētījumiem.

Tika izveidots tiešsaistes interfeiss (4. attēls) un lejupielādēta 796 jaunuzņēmumu galveno lapu datubāze no gaisa transporta jomas.

The screenshot shows a web application interface for annotation. On the left is a dark sidebar with navigation options: Log In, Survey, Training 1 (highlighted), Training 2, and Annotation. The main content area is titled 'INTRODUCTION (MANAGEMENT)' and includes a section for 'number of annotations during this session'. Below this, there is a 'Look at startup SkyTrain' section with a text input field and a 'like' button. A message asks users to skip analyzing startups that don't fit the criteria and provides three checkboxes: 'value is very difficult to identify', 'not a startup', and 'not in the air transportation industry'. There is also a text field for 'other reason (please specify)' and a 'SUBMIT' button at the bottom.

4. attēls. Tiešsaistes saskarne 1. posmam - galveno lapu anotācija

96 anotatoriem ar dažādu izcelsmi, bez profesionālas pieredzes vērtību identificēšanā un angļu valodas zināšanām B2+ līmenī saskaņā ar Kopējo Eiropas pamatprincipu ietvarstruktūru (CEFR) tika uzdots anotēt vērtību tekstus.

Iepriekšējās aptaujas mērķis ir izprast neapmācīto dalībnieku zināšanu līmeni un noteikt vērtību piedāvājuma apjomu un struktūru pirms izglītības procesa. Studentiem tika stāstīts par vērtību piedāvājuma pamatjēdzienu. Pēc tam viņiem tika lūgts uzskaitīt pēc iespējas vairāk vērtību, ko nodrošina gaisa transporta jaunuzņēmumi un uzņēmumi.

Pēc tam, kad dalībnieki tika iepazīstināti ar sākuma koncepciju, viņi uzzināja esošās Vērtības piedāvājuma galvenās teorijas. Pēc tam viņiem tika piedāvātas divas apmācības. 1. apmācībā ir parādīti pieci salīdzinoši vienkārši piemēri, kā noteikt vērtības piedāvājumu jaunuzņēmumu galvenajās lapās. 2. apmācība piedāvāja piecas galvenās lapas, kurās Vērtības piedāvājumu nevar viegli identificēt tā neskaidrās izteiksmes vai nestandarta veidu dēļ, piemēram, video.

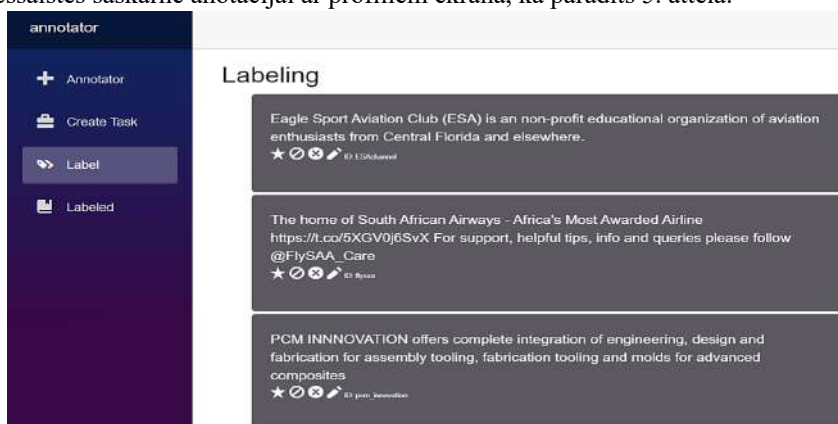
Dalībniekiem ir iespējas identificēt galveno lapu kā tādu, kas nav jaunuzņēmums, var norādīt, ka vērtību ir ļoti grūti identificēt, un var nosaukt lapu kā tādu, kas nav no gaisa transporta nozares. Turklāt viņi tiek aicināti noklikšķināt uz pogas *Patīk*, ja viņi uzskata, ka šis jaunuzņēmums skaidri parāda vērtības piedāvājumu.

Ekspierimenta rezultātā ir radīti bagātīgi semantiskie metadati. Tika apkopotas divas datubāzes: vērtību kopums no atklāta apsekojuma un anotētas vērtības no AT jaunuzņēmumu tekstiem. Anotācijas kvalitāti atkārtoti novērtēja nozares eksperti.

Anotētāju demogrāfija

1. posms Lai apkopotu vispārīgākus novērojumus par gaisa transporta vērtības piedāvājumu, 1. posmam tika pieņemti darbā 96 transporta universitātes studenti ar saistītu, bet joprojām daudzveidīgu specializāciju un/vai profesionālo lomu klāstu transporta jomā: Aviācijas vadība, IT, loģistika. Visi dalībnieki bija vecumā no 20 līdz 42 gadiem.

2. posms Ekspierimenta otrais posms sastāv no Twitter profilu vērtību anotācijas. Twitter profilos ir pārstāvēti 796 jaunuzņēmumi, kuru galvenās lapas tika anotētas eksperimenta 1. posmā. Tika nodarbināti divi pieredzes bagāti anotatori ar dziļām zināšanām par vērtību piedāvājuma koncepciju. Tika izveidota atsevišķa tiešsaistes saskarne anotācijai ar profiliem ekrānā, kā parādīts 5. attēlā.



5. attēls. Tiešsaistes saskarne 2. posmam — Twitter profilu anotācija

Rezultātā tika savākts vērtību korpuss no gaisa transporta jaunuzņēmumu Twitter profiliem.

Taksonomijas izstrādes pieeja

Pamatojoties uz saņemtajiem datiem, ir izstrādātas divas taksonomijas: Gaisa transporta vērtību taksonomija un Vērtību semantiskā taksonomija aviācijā.

Gaisa transporta vērtību taksonomija tika izstrādāta, pamatojoties uz pirmsaptaujas datiem. Semantiskā vērtību taksonomija aviācijā (STVA) tika izstrādāta jaunuzņēmumu galveno lapu anotācijas datos.

Vērtību semantiskā taksonomija aviācijā arī tika saskaņota ar Starptautiskās standartizācijas organizācijas noteikumiem un tika apstiprināta, mijiedarbojoties ar gaisa transporta ekspertiem: vienu starptautiskās aviokompānijas IT direktoru un diviem lidostu komercdirektoriem.

Pamatojoties uz pieeju un izmantotajām metodēm, tika izstrādātas divas vērtību taksonomijas gaisa pārvadājumos. Tālāk autors iepazīstina ar specifisko semantiskās taksonomijas izstrādes un NLP modeļa ieviešanas procesā pielietoto datu analīzes principu – *Viena vārda principu*. Kā daļu no pašreizējā pētījuma, kura mērķis ir automatizēta vērtību atpazīšana, autors pieņem, ka lielāko daļu vērtību varētu uztvert ar vienu vārdu.

Pieņēmuma pamatotība tiek pētīta ar eksperimentiem aviācijas jomā: Tweets anotācijas analīze liecina, ka vairāk nekā 50% no visām identificētajām vērtībām var uztvert ar vienu vārdu: 73% vērtību tika anotētas kā viena vārda vērtība.

12.4. Galvenie atklājumi un izpētes rezultātu īstenošana

Iegūtie atklājumi ir šādi.

- 1) *Konceptuālā analīze un jēdzienu identificēšana* (building a conceptual framework) tika veikta:
 - precizējot izmaiņas pētniecības fokusā uz aviācijas angļu valodu jomām 2002.-2016.gadam un 2017.-2021.gadam.
 - identificējot gaisa transporta vērtību maiņu vēsturē.
 - izstrādājot termina *vērtības piedāvājums* etimoloģisko ģenēzi.
 - tika identificēti galvenie faktori, kas ietekmē vērtības noteikšanu aviācijā.
- 2) *vērtību kopums gaisa pārvadājumos, kas sastāv no 3 daļām* (aptauja, jaunuzņēmumu anotācija un Twitter profilu anotācija) tika uzbūvēts un profesionāli izveidots neironu tīkla modelis
- 3) tika konceptualizēta *viena vārda pieeja* vērtību tveršanai tekstos
- 4) tika izveidotas *divas vērtību klasifikācijas* gaisa transporta jomā:
 - Gaisa transporta vērtības taksonomija, klientu perspektīva
 - Aviācijas vērtību semantiskā taksonomija, nozares perspektīva
- 5) *izstrādāta metodoloģija automātiskai vērtību identificēšanai no tekstiem*

Konceptuālā analīze un koncepciju identificēšana

Lai izveidotu konceptuālu *vērtības piedāvājuma* jēdzienam un *vērtībām* tā ietvaros, noteiktu vērtību lomu gaisa transporta stratēģiskajā vadībā un pārskatītu literatūru par angļu valodas lietojuma izpēti gaisa transportā, tika veikti vairāki pētījumi ar šajā nodaļā aprakstītajiem rezultātiem.

Konceptuālās analīzes un jēdzienu identificēšanas rezultāti ir atspoguļoti šī zinātniskā darba teorētiskajā daļā *Vērtības kā daļa no vērtības piedāvājuma gaisa transporta nozarē*:

- Precizētas izmaiņas pētniecības fokusā aviācijas angļu valodas jomām 2002.-2016.gadam un 2017.-2021.gadam
- Tika identificēta gaisa transporta vērtību maiņa vēsturē
- Tika veikta termina vērtības piedāvājums etimoloģiskā ģenēzē

- Tika identificēti galvenie faktori, kas ietekmē vērtības piedāvājumu aviācijā.

Vērtību korpuss gaisa pārvadājumos

Korpuss

Tiek izveidots vērtību korpuss gaisa transporta jomā, kas sastāv no trim daļām:

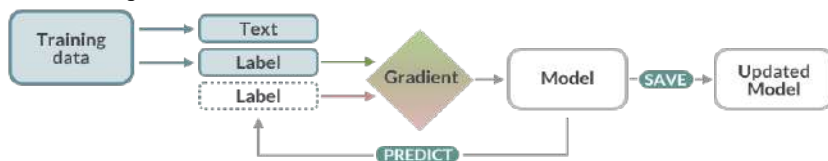
1. Aptauja balstīta un izriet no pašformulētām atbildēm uz atvērto jautājumu: Kādas vērtības sagaidāt no gaisa transporta nozares? ar 241 vērtību.
2. Anotācijas balstītas un izriet no jaunuzņēmumu galveno lapu marķējuma ar 588 vērtībām.
3. Anotācijas balstītas un izriet no jaunuzņēmumu marķējuma Twitter profiliem ar 763 vērtībām.

Īstenošana ar neironu tīklu

Balstoties uz eksperimentu, 2. posms.

Tika piemērota *daļēji uzraudzīta vērtības ekstrahēšanas metode*, veicot šādas darbības:

1. Dokumentu kopums (564 aviācijas jaunuzņēmumu Twitter profili) tika izmantots kā avotu bāze vērtības atpazīšanai
2. Ekspertu marķējums (vērtību anotācija) pieņēma vārdus, kas satur vērtības kā daļu no vērtības piedāvājuma, kas jāmarķē kā vērtības vienības. Tika izmantota kombinētā semantiskās un binārās anotācijas metode.
3. Anotācijas rezultātā beidzot tika iegūta 763 vērtību datu kopa.
4. Priekšapstrādē (filtrēšana, normalizēšana: Python) tika iekļauts Viena vārda princips: vienu vērtību izsaka viens vārds.
5. Uzdevuma vienkāršošana kā binārā klasifikācija: binārā klasifikācija (vērtība/nav vērtība), tā vietā lai katra vērtība būtu klase. ļauj izvēlēties jaunas vērtības
6. Kā binārais klasifikators tiek izmantots neironu tīkla modelis no spaCy bibliotēkas.
7. Apmācība un novērtēšana: 6. attēlā parādīts spaCy modeļa profesionālas ieviešanas process.



6. attēls. Modeļa apmācības process

Izmantotais vārdu korpuss ir "spacy.Corpus.v1" angļu valodai.

Modeļa arhitektūra ir Spacy.TransitionBasedParser.v2. Uz pāreju balstīta parsēšana ir piceja strukturētai prognozēšanai, kurā struktūras prognozēšanas uzdevums ir saistīts ar stāvokļu pāreju sēriju.

Arhitektūra:

Slēptā slāņa izmērs = 64 neironi. Parsētājs sastāv no 3 apakštīkliem:

1. tok2vec slānis (pārveido vārdus (žetonus) vektoros)

2. "apakšējais" slānis (neironu tīkla "pamats", pārvērš vektorus iekšējā attēlojumā)
3. "augšējā" kārtā (plūsmas pārsūtīšanas slānis, kas aprēķina rezultātus no iepriekšējo slāņu vektoriem)

Pēc tam iegūtie punkti tiek izmantoti, lai paredzētu marķējumu vārdam.

7. attēlā parādīts apmācības progress, kur ents_p, ents_r, ents_f ir NER uzdevuma precizitāte, atsaukums un f-vērtējums.

===== Training pipeline =====

[i] Pipeline: ['tok2vec', 'ner']

[i] Initial learn rate: 0.001

E	#	LOSS TOK2VEC	LOSS NER	ENTS_F	ENTS_P	ENTS_R	SCORE
0	0	0.00	75.80	55.76	49.89	63.20	0.56
0	10	0.99	710.78	71.61	61.04	86.59	0.72
0	20	3.28	451.60	72.88	62.92	86.59	0.73
0	30	3.09	400.33	76.67	68.96	86.32	0.77
0	40	5.55	289.59	81.32	78.57	84.27	0.81
0	50	9.35	267.72	82.42	80.03	84.95	0.82
0	60	11.34	254.40	82.69	79.70	85.91	0.83
0	70	10.72	225.34	83.03	79.97	86.32	0.83
0	80	9.20	236.51	81.70	80.67	82.76	0.82
0	90	6.87	244.33	84.25	81.30	87.41	0.84
0	100	7.37	247.18	84.38	81.08	87.96	0.84

[+] Saved pipeline to output directory

spacy_output\model-last

[2021-11-10 16:17:39,067] [INFO] Set up nlp object from config

[2021-11-10 16:17:39,076] [INFO] Pipeline: ['tok2vec', 'ner']

[2021-11-10 16:17:39,079] [INFO] Created vocabulary

[2021-11-10 16:17:39,081] [INFO] Finished initializing nlp object

[2021-11-10 16:17:39,753] [INFO] Initialized pipeline components: ['tok2vec', 'ner']

8. attēls. Cauruļvads

Sākot no rezultāta 0,56, pēc 10 apmācību iterācijām rezultāts palielināts līdz 0,72, pēc 60 apmācību iterācijām rezultāts palielināts līdz 0,83.

Eksperiments tika apturēts pie 100 iterācijām ar punktu skaitu 0,84 šāda iemesla dēļ:

1. uzlabošanās izlīdzinās (kļūst pakāpeniska) (sk. 3. tabulu.)
2. lai izvairītos no pārmērīgas pielāgošanas

Visu līmeņu taksonomijas grupas nav saistītas ar konkrētām runas daļām, un tās vieno viens un tas pats semantiskais jēdziens.

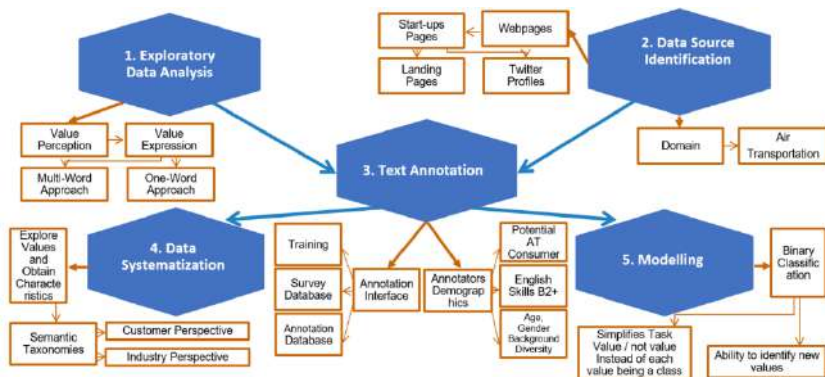
Kategorijas netiek samazinātas, izslēdzot daļējus sinonīmus, jo semantiskās atšķirības saglabāšana var būt svarīga mārketinga speciālistiem un klientu apkalpošanai. Taksonomiskās klases un apakšklases tika izstrādātas no 196 vērtībām, kas apraksta aktivitātes, procedūras, objektus, emocionālo pieredzi un ņemot vērā kontekstu, kādā tās varētu būt saistītas ar gaisa transportu.

Analizējot terminus un kontekstu, ir izveidotas septiņas taksonomiskās klases un sešas apakšklases. Katrs no 196 zemākā līmeņa terminiem iekļaujas vismaz vienā taksonomiskajā grupā. Tomēr daudzi no tiem pieder vairāk nekā vienai taksonomijas grupai un tika piešķirti noteiktai grupai saskaņā ar piecu ekspertu vienošanos: diviem no ekspertiem ir grāds Lingvistikā un vairāk kā 20 gadu pieredze lingvistiskajā pētniecībā, vienam ekspertam ir doktora grāds un vairāk kā 20 gadu pieredze Dabiskās valodas apstrādē, viens eksperts ir reģionālās lidostas vadošais menedžeris, kuram ir doktora grāds, bet viens eksperts ir Starptautiskās aviokompānijas Informācijas tehnoloģiju direktors ar vairāk kā 20 gadu pieredzi gaisa pārvadājumos. Kategorijas nav viena otru izslēdzošas, taču šīs taksonomijas mērķis nav nodrošināt pārklājošu semantisko klasterizāciju.

Katra grupa un katrs termins parāda vairākus cilvēkus, kas nosauc šo vērtību. Visbiežāk lietotie termini (izņemot pieturas vārdus) ir: Augstas (215), Izmaksas (160) un Vienkāršs (53). Katras grupas svars ir tās sastāvdaļu summa. Lielākā grupa ir Atribūti ar 2455 un mazākā ir Statuss/Nodarbošanās (Cilvēki) ar 203..

Piedāvātā metodika automatizētai vērtību identificēšanai no tekstiem

9. attēlā parādīts pētījuma gala rezultāts un pētījuma mērķa sasniegšana – Vērtību automatizētas identificēšanas no tekstiem metodika.



9. attēls. Piedāvātā metodika automātiskai vērtību identificēšanai gaisa transporta jomā (autora veidota)

Ierosināto metodoloģiju ierobežo viena joma – Gaisa transports kā Gaisa transporta joma tika akceptēts un testēts gadījuma pētījumā šajā pētījumā, un izstrādāto Dabiskās valodas apstrādes modeli var īstenot Gaisa transporta tekstiem.

Turklāt metodoloģija ir piemērojama jebkurai jomai un, ja konsekventi veic visus pasākumus, jo semantiskās anotācijas un datu kategorizēšanas metodēm nav jomu ierobežojumu un jaunuzņēmumi kā datu avots aptver lielāko daļu nozaru.

12.5. Pētniecības rezultātu diskusija

Analīze, kas sniegta šajā disertācijā, atbalsta sākotnējo tēzi, ka vērtības kā daļu no vērtības piedāvājuma var identificēt automātiski un palielināt vērtību skaitu. Turklāt šī pētījuma mērķis ir uzlabot esošo praksi un literatūru gan empīriski, gan teorētiski.

Teorētiskais un empīriskais ieguldījums

Šīs analīzes palielina teorētisko izpratni par vērtības jēdzienu, klientu izpratni par vērtību un teorētisko sistēmu vērtības automatizētai identificēšanai un paplašināšanai.

Šī izpēte ir pirmā, kas veic automatizētas vērtības noteikšanas un paplašināšanas uzdevumu. Tas nodrošina teorētisko sistēmu automatizētai vērtības identifikācijai un paplašināšanai, ko var piemērot jebkurā jomā.

Šis pētījums ir pirmais, kurā vērtības tiek uzskatītas par daļu no vērtības piedāvājuma kā primārās izmeklēšanas jomas. Turklāt šajā pētījumā izmantota jauna pieeja, lai pieņemtu jaunuzņēmumu tekstus kā pētniecības bāzi vērtību identifikēšanai un paplašināšanai. Tādējādi tiek paplašinātas pastāvošās metodes uzņēmuma vērtības piedāvājuma izstrādei. Turklāt iegūtie dati liecina, ka vērtības var būtiski paplašināt un kategorizēt, lai atbalstītu lēmumu pieņemšanas procesu. Tādējādi tika konstatēts, ka dziļās mācīšanās rīkus var izmantot, lai automatizētu vērtības noteikšanas un paplašināšanas procesu.

Praktiskā nozīme un nozares apskati

Trīs gaisa transporta nozares eksperti ar augstu zināšanu līmeni - diviem starptautisko lidostu komercdirektoriem un vienam lielākās aviokompānijas IT vadītājam - tika lūgts salīdzināt izstrādāto Vērtību piedāvājuma taksonomiju ar IATA Airline taksonomiju. Viņi dalījās arī ar savām idejām par to, kā taksonomijas var noderēt nozarei un COVID-19 ir pārsteidzoši maza nozīme no respondentu puses 2020. gada aprīlī. Tālāk ir sniegts ekspertu pārskata kopsavilkums:

Gaisa transporta nozares ekspertu atsauksmes:

International Hub Airport komercdirektors: “Tirgotājus ļoti interesē negaidīti atradumi un vērtību formulējumi, kas iepriekš nebija redzami. Labu perspektīvu var redzēt, izmantojot aviācijas vērtību taksonomijas projektu vadībā.”

Aviokompānijas IT vadītājs: “Šīs taksonomijas ir ārkārtīgi noderīgas šādam mērķim: lai izprastu pakalpojumu/produktu būtību un sistematizāciju. Mūsu uzņēmums ir izmantojis unikālu semantisko vērtību taksonomiju aviācijā, izstrādājot mūsu jauno tiešsaistes pārbaudes un iepirkšanās sistēmu.”

Reģionālās starptautiskās lidostas komercdirektors: “Automatizācija un vērtību sistematizācija gaisa transportā ļauj ātrāk pieņemt lēmumus par stratēģiskiem jautājumiem, jo tā parāda jaunas vērtības un vērtību ķēdes pārveidošanu; taksonomijas ļauj formulēt galvenos darbības rādītājus pa virzieniem, sadalīt atbildības jomas starp visiem pievienotās vērtības ķēdes posmiem.”

Secinājumi un ieteikumi

Secinājumi

Šis pētījums ir izstrādāts, lai palīdzētu uzņēmumiem izmantot datu apstrādes potenciālu, lai radītu jaunas uz datiem balstītas lēmumu atbalsta sistēmas vērtības piedāvājuma izstrādes procesā. **Pētījuma mērķis ir izstrādāt metodiku automatiskai vērtību identifikācijai un paplašināšanai gaisa transporta jomā** ir sasniegts, un ir izstrādāta metodika.

Galvenais pētījuma jautājums “*Vai var identificēt vērtības un palielināt vērtību skaitu automatiski?*” ir atbildēts apstiprinoši.

Piecās šī pētnieciskā darbanodaļās ir izstrādāta jauna vērtības piedāvājuma metodoloģija kā daļa no vērtības piedāvājuma un paplašināšanas, kā arī modelis un parādīts, kā tās ļauj iegūt konkrētus rezultātus vērtības piedāvājuma un paplašināšanas procesa automatizēšanā.

Lai pierādītu vērtību automatiskās identifikācijas un paplašināšanas lietojamību, ir apspriestas mūsdienu perspektīvas uz datiem balstītu biznesa procesu izvēršanā. Ir pierādītas priekšrocības, ko sniedz datu analīzes izmantošana, izmantojot Dabiskās valodas apstrādes pieeju un teksta ieguves metodes, lai uzlabotu uzņēmējdarbības veiktspēju: ietaupot laiku un pūles. Ir sniegtas vadošo rūpniecisko analītiķu atsauksmes, tostarp ieskats par digitālās stratēģijas priekšrocībām uzņēmējdarbības procesos un uz datiem balstītu lēmumu pieņemšanas atbalsta sistēmu ieviešanā: vērtību automatizācija un sistematizācija gaisa transportā ļauj ātrāk pieņemt lēmumus par stratēģiskiem jautājumiem.

Jaunu un unikālu vērtību nozīmi ilustrē šādi secinājumi:

- Ņemot vērā pieaugošo konkurenci un gaisa transporta liberalizāciju, pieaug unikāla un jauna vērtības piedāvājuma nozīme aviācijā, un tas strauji kļūst par stratēģisku prioritāti.
- Tradicionālās vērtības gaisa pārvadājumos ir cena, kvalitāte un ātrums. Pēdējās desmitgadēs gaisa transporta komercializācija izraisa konkurenci un pēc tam jaunu vērtību meklēšanu.
- Vērtības radīšana ir viens no galvenajiem uzņēmuma mērķiem, un pēc nesenās liberalizācijas gaisa transporta nozare saskaras ar neizpildītu vajadzību noteikt jaunu vērtības piedāvājumu. Šī procesa automatizācija būtu izdevīga.
- Lai automatizētu vērtības atpazīšanas procesu, jaunuzņēmumi uzrāda pieklājīgu datu avotu vērtības piedāvājuma analīzei, jo to mērķis ir strauji augt un mērogoties, informācija galvenajās lapās ir atklāta un faktiskā.

Automatiskās vērtības noteikšanas galīgā metodoloģija ietver šādus soļus: izpētes datu analīze, datu avota identifikācija, teksta anotācija, datu sistematizācija un modelēšana.

Autors šo darbu uzskata par koncepcijas apliecinājumu, parādot, ka ir vieta argumentācijai par automatizāciju vērtību identifikācijai un paplašināšanai un modeļa izstrādei. Šā instrumenta tūlītēja piemērošana ir paredzēta gaisa transporta

uzņēmumos, kas izstrādā jaunu vērtības piedāvājumu pašreizējā problemātiskajā labklājībā un politiskajā vidē.

Turpmākos pētījumus var veikt sekojošajam:

1. dažādu jomu iekļaušana vērtību korpusā
2. iegūtās metodoloģijas pielietošana noteiktos laika periodos, lai iegūtu vērtību laikrindas, lai paplašinātu vērtību metriku un sekotu vērtību maiņai
3. vērtību noteikšana no oficiālajiem noteikumiem varētu sniegt zināmu ieskatu par oficiālo politiku no dažādiem viedokļiem, salīdzinot ar nozares viedokļiem.

Pētījuma reproducējamību var pierādīt, sekojot metodoloģijas soļiem un iegūstot augstu ticamības pakāpi, novērtējot rezultātus un precizitāti.

Rekomendācijas

Pamatojoties uz iegūtajiem rezultātiem, diskutējām ar ekspertiem un secinājumiem, šī pētījuma autore ir formulējusi sekojošus ieteikumus:

Lēmumu pieņēmējiem gaisa transporta jomā: automātiska vērtību noteikšana kā daļa no vērtības piedāvājuma ir izdevīga, lai ievērojami samazinātu laiku un pūles, lai izprastu esošās vērtības tirgū un atklātu jaunas pieprasījuma tendences.

Mārketinga speciālistiem gaisa transporta jomā: respondentu atbilžu formulējums uz atvērtiem jautājumiem par gaidāmajām vērtībām no aviācijas atbalsta, lai saprastu, ko un kā patērētājs sagaida no gaisa transporta nozares.

Nozares analītiķiem: automātiskā vērtēšana nodrošina zemāku izmaksu analītisko ieskatu tirgū vērtības piedāvājuma, stratēģiskās plānošanas un jaunu produktu tendenču kontekstā.

**THE PROMOTION WORK PRESENTED TO THE TRANSPORT
AND TELECOMMUNICATION INSTITUTE TO OBTAIN THE
SCIENTIFIC DEGREE – DOCTOR OF SCIENCE IN
ENGINEERING**

OFFICIAL REVIEWERS:

Dr.sc.ing. Irina Jackiva
Professor, Transport and Telecommunication Institute, Riga, Latvia

Dr. Eftihia Nathanail
Professor of Transportation Engineering, University of Thessaly,
Greece

Dr. Vera Sheinman
NLP Software Engineer, Apple Inc., Cambridge, United Kingdom

The defence of the thesis will be held on December 21, 2022, 3p.m. at the Promotion Council of the Transport and Telecommunication Institute. The address is 1 Lomonosova Street, conference hall 130, Riga, Latvia, tel. +371 67100594, fax: +371 67100535.

CONFIRMATION

I hereby confirm that I developed the promotion work that is presented to the promotion council of Transport and Telecommunication Institute to obtain the scientific degree of Doctor of Science in Engineering. The promotion work has not ever been presented to any other promotional council to the scientific degree.

O. Zervina

The promotion work is written in English. It consists of 5 chapters, the introduction and conclusions, as well as 14 appendices. It contains 194 pages, 39 figures, and 24 tables. The list of references contains 301 titles.

CONTENT

ABSTRACT	32
1. RELEVANCE AND MOTIVATION OF THE RESEARCH	33
2. QUESTIONS AND TASKS OF THE RESEARCH	34
3. AIM, OBJECT AND SUBJECT OF THE RESEARCH	34
4. SCOPE OF THE RESEARCH.....	34
5. THE APPROACH.....	36
6. METHODOLOGY AND THE METHODS OF THE RESEARCH	36
7. SCIENTIFIC NOVELTY OF THE RESEARCH.....	36
8. THEORETICAL AND PRACTICAL VALUE OF THE RESEARCH	37
9. APPROBATION OF THE RESEARCH	37
10. STRUCTURE OF THE THESIS	38
11. STATEMENTS PUT FORWARD FOR DEFENSE.....	38
12. SUMMARY OF THESIS CHAPTERS	39
12.1. Values as Part of Value Proposition in Air Transportation	39
12.2. Methodological Approaches: State of the Art	41
12.3. Methodology and Research Design	44
12.4. Key Findings and Research Results Implementation	49
12.5. Research Results Discussion	53
Conclusions and Recommendations	54
Autora publikāciju bibliogrāfiskais saraksts/ Publications with author's participation.....	56
Literatūras saraksts/References.....	57

ABSTRACT

The Thesis “A Methodology of Automated Identification of Values in the Air Transportation Domain on the Base of E-Texts” is written by Olga Zervina under the supervision of Dr.sc.admin, professor Yulia Stukalina, Dr.sc.ing., professor Dmitry Pavlyuk, and Dr. Neil Rubens.

As an object of research in this work, the value proposition in the air transport domain is determined. The work is focused on the development of a data-driven decision-support model, identifying the tools and procedures that decision-makers might use to determine new values as part of value proposition.

The requirement for new values and for increasing the number of values in Air Transport is vividly demonstrated in 2018 in the International Air Transport Association (IATA) Industry Affairs Committee (IAC) Report “Future of the Airline Industry 2035”. Recommendation 12 of the document clearly states that “as new technologies and value shifts change how and why people work, the industry will need to invest in skills for future aviation leaders and workers and communicate the benefits of working in the sector”.

The argument for the wider value of aviation is reflected in IATA’s Industry Affairs Committee Report as the combination of trends in Big Data, automation and the internet of things. These changes are expected to lead to new opportunities, transforming how people and devices are monitored and marketed to in real-time. It is beneficial for air transport companies to take reward of advances in automation, new transport modes, and consumer attitudes.

This research supports the concept that controlling data on value will have an advantage over existing competitors in developing new niches. The novelty and aim of the research lies in the state-of-the-art methodology that allows text-mining techniques as part of Natural Language Processing to be applied to the numerous start-up landing pages.

Value Identification is a complex and difficult problem. The author uses an NLP approach to identify methods and features well suited for this problem. A data-driven technique is taken, i.e., the author first constructs the dataset and then analyzes it using computational linguistic methods to identify which features and methods perform the best. The author also takes the corpus linguistic approach, i.e., reliable language analysis is more feasible with corpora collected in the field in its natural context.

Main results of this thesis are presented at 14 international scientific conferences and published in 12 scientific papers. The thesis consists of 5 chapters and includes 194 pages, 39 figures, 24 tables in the main body, 14 appendixes and 301 references in the bibliography.

1. RELEVANCE AND MOTIVATION OF THE RESEARCH

Value creation is one of the primary goals of any business entity (Prasada et al, 2021). The strategy of a company starts with the developing its *value proposition* as (Lanning and Michaels, 1988) named a simple statement that summarizes why a customer would choose your product or service: *our affordable plane tickets, our comfortable business-class lounge*. If several people undertake to construct a company's value proposition, their results may differ in focus and subject matter. And the question is, which of these value propositions is correct? Perhaps they all do not reflect the values of the market? To find out, companies need to conduct costly and long-term marketing research that includes a fair amount of human error. To improve the situation, a more objective approach must be found, possibly with technical support, which is quite possible with the potential of the modern world (Ebel et al., 2016).

Automation of value identification would definitely lead to the extension of number of values. Moreover, Kaplan and Norton (2008) advise a company's strategy to be grounded "on a differentiated customer value proposition". Recent research shows the tremendous difference between the performance of companies depending on the number of values offered to the customers. When buyers assess a service or a product, they compare the perceived value of a product or a service to the price. (Almquist, Senior and Bloch, 2016) consider that for marketers, understanding and analyzing how customers think about the pricing side of the equation is a known and straightforward process. For strategic managers, understanding what consumers actually value and why they appreciate it, on the other hand, is far more difficult. Finally, business performance is improved for organizations that provide many components of value (Almquist, Senior and Bloch, 2016).

The capacity to recognize values at an early stage is critical for business performance. This research focuses on Air Transportation industry, which has recently experienced deep upheavals due to pandemic and geopolitical unrest leading to new values requirements.

The nature of value constantly depends on the viewer's perspective (Almquist, Cleghorn and Sherer, 2018). Competition is won by those companies who create the most diverse, unique, novel and demanded value proposition. Thus, businesses are faced with the challenge of identifying and expanding the value proposition.

To accomplish this task, companies use customer surveys and competitive analysis to define their value proposition. However, identifying values manually is a time-consuming task; and the number of obtained values is often limited.

In today's age of information technology and Big Data, it is imperative to employ these technology and massive amount of data in the process of developing a value proposition by companies. There exists data on values in companies' descriptions of their products on homepages, marketing materials and social network profiles. Major message indicating values is formulated by texts. Analysing these data using text mining techniques provides the opportunities for value identification

and expansion (increasing the number of values). Automated value recognition tools can be regarded as a compelling additional challenge in the business context.

2. QUESTIONS AND TASKS OF THE RESEARCH

To solve the problem of automation of value identification in the e-texts of air transportation domain, the following *research questions* can be formulated:

1. How the notion of value is defined in literature?
2. Why does aviation industry require value expansion?
3. Can values be identified manually?
4. Can values be categorized?
5. Can number of values be identified and expanded automatically?

In order to answer these research questions, the following *research tasks* arise:

1. To investigate literature on the philosophical ground and factual implication of value concept
2. To explore aviation industry values
3. To identify air transport industry specific factors that influence the development of values
4. To conduct an experiment annotating texts labelling values in the aviation context
5. To develop value categorization in the form of taxonomies
6. To discover additional values in comparison with the traditional number of values in aviation organization
7. To create an automated methodology for value identification.

3. AIM, OBJECT AND SUBJECT OF THE RESEARCH

The research aim is to develop a methodology for automatic identification and expansion of values in the domain of air transportation.

Object of the research is value proposition in the air transport domain.

Subject of the research is a methodology of a Natural Language Processing analysis allowing to detect values as a part of value proposition.

The interdisciplinary approach to the object and subject of the research, the strategy of which is aimed at obtaining new synthetic knowledge about the ways of extracting values, unlocks great prospects for conducting comparative scientific research.

4. SCOPE OF THE RESEARCH

To explore the possibility of automatic value identification and expansion, this research narrows down the domain to the air transportation industry.

The requirement for new values and for increasing the number of values (from here on in this thesis *increasing the number of values* refers to as *value expansion*) in Air Transportation is vividly demonstrated in the International Air Transport Association (IATA) Industry Affairs Committee (IAC) Report “Future of the Airline Industry 2035” (IATA, 2018). Recommendation 12 of the document

clearly states that “as new technologies and value shifts change how and why people work, the industry will need to invest in skills for future aviation leaders and workers and communicate the benefits of working in the sector”.

The argument for the wider value of aviation is reflected in IATA’s IAC Report (IATA, 2018) as the combination of trends in Big Data, automation and the internet of things. These changes are expected to lead to new opportunities, transforming how people and devices are monitored and marketed to in real-time. There is the benefit for air transport companies to take reward of advances in automation, new transport modes, and consumer attitudes.

This research supports the concept that controlling data on value will have an advantage over existing competitors in developing new niches. As (Almquist, Senior and Bloch, 2016) proved, companies that scored high on four or more elements of value had recent revenue growth four times greater than that of companies with only one high score.

Companies attempt to develop their value proposition via comparison to rivals and conducting market research. To complete them manually, these procedures are labor-intensive and time-consuming. Automation has a distinct benefit in that it helps companies to detect more values while also suggesting new, unique, and previously unconsidered possibilities for businesses. Äyväri and Jyrämä, (2017) examined three popular tools for building value propositions: Value Proposition Canvas by Osterwalder and Pigneur (2010), The Value Proposition Builder™ by Barnes and colleagues (2009), and People Value Canvas by Wildevuur and colleagues (2015). All of them are multi-stages tools involving market and customer analyses.

Thus, automation has a distinct benefit in that it assists companies in detecting more values while also suggesting new, unique, and previously unconsidered possibilities for businesses. In today’s online age, companies mostly publish value proposition using their webpages including landing pages, online advertisements and social network profiles. Value proposition is delivered by means of text, visual and sound techniques. This research focuses on textual value proposition and employs NLP understand human language by analyzing text. Text mining is used to extract information from unstructured content.

The automation of value proposition in the research involves texts from landing pages and Twitter profiles of start-up companies in the air transport domain. Start-ups are chosen as a source of data of value as part of value proposition under the following criteria: start-ups are among the first companies to indicate new and unique values as their value proposition is based on a novel business model or innovation. Besides that, start-ups provide their information on novel value openly and willfully as one of the start-ups’ goals is to continuously attract funding. The permanent process of search for funding reflects the nature of start-ups: constant growth aimed globally. To attract attention and offer convenient way to interact, start-ups provide information about themselves in the form of landing pages to unified databases, collecting thousands of start-ups.

5. THE APPROACH

Value Identification (VI) is a complex and difficult problem. The author uses an NLP approach to identify methods and features well suited for this problem. A data-driven technique is taken, i.e., the author first constructs the dataset and then analyzes it using computational linguistic methods to identify which features and methods perform the best. The author also takes the corpus linguistic approach, i.e., reliable language analysis is more feasible with corpora collected in the field in its natural context.

6. METHODOLOGY AND THE METHODS OF THE RESEARCH

This study uses qualitative and quantitative methods to create a methodology for the automatic identification of values in the field of air transport. First step is the theoretical analysis on values aimed at determining meta-characteristics of the process of value identification and expansion in the air transportation domain. It includes secondary data analysis: review of theoretical literature, previous research, policy documents and theoretical framing as evaluation and conceptualization of approaches and notions. Second step was to develop the research procedure, so modern approaches to information extraction and systematization were applied: semantic annotation combined with a binary classification approach, open-ended survey, categorization and organization into taxonomies. Implementation of the results was tested using a classifier: a neural network.

7. SCIENTIFIC NOVELTY OF THE RESEARCH

The novelty of the research lies in the state-of-the-art methodology that allows text-mining techniques as part of Natural Language Processing (NLP) to be applied to the numerous start-ups landing pages to collect data as the part of automation process of value identification and extension, detailing:

1. Novel task was proposed - value identification from texts, and a new Named Entity is introduced: Value Entity Recognition (VER).
2. Methodology for automation of value identification from texts with the goal to increase the number of values was developed.
3. Corpus/dataset of values in air transport was created.
4. Systematization of values: two value categorizations were developed in the form of taxonomies:
 - Air Transport Value Taxonomy, customer perspective.
 - Semantic Taxonomy of Values in Aviation, industry perspective.
5. A conceptual framework was built:
 - Value expression: one-word concept developed.
 - Social aspects of value: individual vs group values ratio concept developed.
 - Air transport value shift through history was identified.
 - Etymological genesis of the term value proposition developed.
 - Key factors influencing value proposition in aviation identified.

8. THEORETICAL AND PRACTICAL VALUE OF THE RESEARCH

This research constitutes the first study considering quantitative characteristics of values as a part of value proposition as the primary field of investigation. Additionally, this study uses a novel approach for taking texts of start-up companies as a research base for identifying and expanding values. This broadens the existing techniques on developing value proposition of a company. Moreover, the obtained data show that values can be significantly expanded and categorized to support the decision-making process. Through this, it was identified that deep learning tools can be utilized to automate the process of value identification and expansion.

Theoretical contribution of the research can be formulated as follow:

1. This research provides the novel concept of a possibility for automated value identification and expansion which leads to the potential categorization.
2. This research is the first to conduct a task of automated value identification and expansion. It provides the theoretical framework of automated value identification and expansion that can be applied to any domain.
3. The research explains which factors should be considered in automated identification-and expansion of values as part of value proposition: notion of value, inter-annotator agreement, one-word approach in the classification, classifier's accuracy.
4. Finally, the provided categorizations of values in the form of taxonomies contributed into the following: a common language for sharing concepts, knowledge discovery (patterns, trends), allowing for reuse of information rather than recreation, match items or contents sharing the same concepts, provide metadata for identification, comparison, analysis.

The following statements highlight the practical importance of the research:

1. Strategic decision-making scope is widened by providing extended choice of values based on innovations and novel business models.
2. Taxonomies allow elaborating common language for new concepts; provide consistent data for identification, comparison, and analysis.
3. Automation provides opportunities for identification industry value shift through time and new values can be documented using time- and resource-saving technique.
4. Novel methodology has a potential applied to different domains enabling identification and extending values beyond air transport.

9. APPROBATION OF THE RESEARCH

The basic concepts and findings were reported and discussed at scientific conferences and seminars. The author has 12 publications on the subject of the thesis and 23 conferences, workshops and real-life business-processes in Latvia, Japan, Italy, Estonia and Lithuania. The developments and findings made within this research were used in pedagogical practice at the Transport and Telecommunication Institute in form of case studies incorporated into the subject "Digital Marketing".

10. STRUCTURE OF THE THESIS

The work consists of the Introduction, 5 chapters, the Conclusion, and 14 appendices. It contains 194 pages, 39 figures and 24 tables. The list of references and information sources contains 301 titles.

The Introduction presents the motivation of the research, the aim, tasks, the object and the subject, the novelty, the empirical and theoretical contribution of the research, as well as its structure.

The first chapter describes the development of aviation strategic management as air transport is taking as a case study in this research and the concept of Value Proposition and values as part of this concept. Values and value proposition in this context are the object of this study. The data source for this research is online texts of start-up companies, their benefits as a source of data are discussed in this chapter.

The second chapter discusses data-driven decision making and decision-support systems, the benefits of making decisions based on data and provides methodological background of the research. The potentials of information extraction methods for supporting business processes are presented. Narrowing the methods of information extraction to named entity recognition and binary classification approaches as some of the main applied methods of text mining of this study is presented. Value Entity Recognition is offered as a resulting combined method.

The third chapter describes the procedure of the experiment, taxonomies development principles and a novel one-word approach to the value identification by annotators and classifiers.

The fourth chapter presents key findings and response to the research questions: conceptual analysis, the corpus of values, hierarchical classifications of values, possibility to automate value identification and expansion process by developing natural language processing model.

The fifth chapter offers empirical and theoretical contribution of the Thesis and discussion on value analysis and practical importance of the results.

The *conclusions* list the main results of the study and highlight the most important findings. It is shown that the statements put forward for defence have been confirmed. Finally, perspective for further research are highlighted and recommendations for the industry analysts and decision-makers are provided.

11. STATEMENTS PUT FORWARD FOR DEFENSE

The following statements are put forward for defense:

1. Values can be identified from texts (landing pages, Twitter profiles)
2. Values in air transport domain can be categorized via taxonomies
3. A methodology allows to automatically identify values from textual information about the company

12. SUMMARY OF THESIS CHAPTERS

12.1. Values as Part of Value Proposition in Air Transportation

Values as a part of Value Proposition concept

New values are being adopted by companies to increase their profitability and resilience. According to (Almquist, Senior and Bloch, 2016) in his elements of value study, success of companies is driven by ability to deliver on multiple Elements of Value. In their research, Almquist et al give examples of the number of values for the world leading smartphone brands: for example, LG delivers 3 elements, Samsung 10 and Apple 12); a company that delivers good performance on bigger number of elements than other market players can be attracting more customers and can charge higher price than competitors do. Examples: Red Bull - the company sponsors extreme sports connecting sports enthusiasts worldwide), Nordstrom (the company acquired Trunk Club giving sense of affiliation and belonging).

Though (Lanning and Michaels, 1988) first used the expression “value proposition” (VP) in a 1988 work document for the consulting company “McKinsey and Co”, (Almquist, Senior and Bloch, 2016) suggested an original strategy based on a differentiated customer value proposition. A suggested set of value was called Elements of Value. Elements of Value where categories are based on the classic Maslow’s Hierarchy of Needs (Maslow, 1943).

Maslow’s Hierarchy of Needs (Maslow, 1943) so far has been serving as a basic for evaluation of human nature and human requirements for the well-being. For the first time, (Almquist, Senior and Bloch, 2016) have adopted this hierarchy of human needs to the hierarchy of the values that are expected from business by customers. This idea serves as a brilliant example of cooperation between business and science.

There is comparatively few research on analyzing value propositions in online start-ups. (Su-C Li, 2007) argues that a properly constructed value proposition is essential to the value creation process in e-business, and value co-production is the building blocks for value protection mechanism in network economy.

(Äyväri and Jyrämä, 2017) in their article “Rethinking value proposition tools for living labs provide a conceptual analysis on value proposition tools to be used in future empirical research and in building managerial insight. Three managerial methods for developing value propositions are examined in the context of the conceptual analysis, which is based on living labs frameworks and current theoretical advances surrounding the notion of value. Among findings in the context of the living labs approach, the Value Proposition Builder seems to conflict with the ideas and premises of user-centric innovation processes (Äyväri and Jyrämä, 2017).

This research adopts the following terms in the context of value as a “statement that clearly identifies what benefits a customer will receive by purchasing a particular product or service” (Rintamäki et al, 2007): Value, Value Proposition, Customer Value, Customer Value Proposition.

Competitive advantage after recent liberalization delivers new challenges; a requirement to identify and provide values for the passengers, as well as delivering internal value, should be taken into consideration by aviation industry. In aviation

traditional dominant values are price, comfort, and speed, and it has been dominating for decades (IATA Passenger Surveys, 1991-2019; Bain&Company, 2011).

Table 1 summarizes factors influencing value proposition in aviation. This table provides the factor, its economic characteristics and specific impact on value.

Table 1. Factors influencing value proposition in aviation (Author)

Factor	Economic characteristics	Impact on value
Perishable product	aviation product is immaterial in nature, rapid-perishable, not stored	in the short-term, expenses to provide capacity are unmet; pressing price discounting
Similar products	the airline product is homogenous in nature, similar across airlines	novel products initiative is quickly imitative by competitors; no differentiation value
Low marginal cost structure	low marginal costs, high fixed costs for additional passengers	dependence on high fixed costs restricts value proposition range
Significant exit barriers	long term aircraft and airport capacity; rigid full-service network carrier business model	fixed unalterable value proposition; any additional ways of utilizing existing resources are beneficial (e.g. by pursuing new values)
Capacity	New aircraft and airport capacity: long term return on investment	capacity is fixed / inflexible problem: underutilization (if insufficient demand) inability to increase capacity easily
Industry growth	volatile and heterogenous	unstable periods cause a conservative approach to value proposition
Heterogeneity of conditions	Airports and air carriers are subject to the specific policies at their domestic market	it impacts the competitive collaboration between the companies in a way not related to value proposition or efficacy; so, no need for new values search
Recent liberation, monopolization	less competition	Value as an afterthought

Start-up company: a source of concentrated and open value proposition

European Commission defines start-ups as “often tech-enabled, in general combine fast growth, high reliance on innovation of product, processes and financing, utmost attention to new technological developments and extensive use of innovative business models, and, often, collaborative platforms” (European Commission, 2016).

Increasingly, nowadays business environment, data-driven innovation and the production of economic value are generated by several organizations rather than by a single organization or by a conventional value chain (Hein et al, 2019). As a replacement for, combinations of different data form multidisciplinary databases, enriching cross-industrial and socio-tech potential of businesses (Hein et al, 2019; Oliveira and Loscio, 2018; Yoo et al, 2010). Multiple start-ups from one industry domain can be considered as an ecosystem that provides among others information on value proposition.

As start-ups’ goal is to grow and scale rapidly (Giardino et al, 2014) with “high reliance on innovation of product, processes” (European Commission, 2016), information on their landing pages is open and actual; it provides a good data source

for analysing existing and prospective value proposition. Figure 1 summarizes the potential of start-up companies to become a source of values for a certain domain.

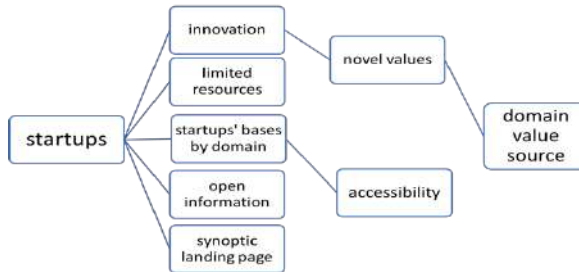


Figure 1. Start-ups as a source for values (author's construction)

Concluding remarks

1. At the corporate level of the development of aviation decisions are made using a tool such as strategic management.
2. Many experts in the field of strategy include creating value as a core thing among the potential components of a successful company's strategy.
3. In the face of growing competition and air transportation liberalization, value proposition in aviation is increasing and fast becoming a strategic priority.
4. The traditional values in air transportation are price, quality and speed. In the recent decades, commercialization of air transportation initiates competition and, subsequently, new values search.
5. Value creation is one of the main company goals, and after the recent liberalization air transportation industry faces the unmet need of identifying new value proposition. Automatization of this process would be beneficial.
6. To automate the process of value recognition, start-up companies present a decent data source for analysing value proposition since their goal is to grow and scale rapidly, information on their landing pages is open and actual.

12.2. Methodological Approaches: State of the Art

Traditional, non-data-driven approaches to decision-making in creating value proposition, such as offered by Osterwalder and Pigneur (2010), typically follow a data-independent, manual approach involving one or few participants.

To model how companies implement their businesses, different Business Models (BM) systems, for example, broadly acknowledged Business Model Canvas including value proposition creation were created supporting a universally accepted interpretation of Business Model in a company. Nevertheless, those typical, non-data-driven decision-making processes experience a number of disadvantages that restrict their utility.

Information Extraction

With the explosive growth of the World Wide Web, admits (Chen, 2014), there is a vast wealth of information in semi-structured and unstructured documents (e.g., web pages and full-text documents) available on the Web. Efforts to automatically access and distil the information from these sources have thus been ongoing for the past decade. This area is known as Information Extraction (IE). In contrast to Information Retrieval (IR), whose task is to find out and rank relevant documents according to queries representing a user's information need, a broad goal of IE systems is to extract structured data from semi-structured and/or unstructured machine-readable documents, where the specific type and structure of the information to be extracted depend on the need of the particular application.

Web IE tools, also referred to as wrapper induction systems, (e.g. Lixto (Baumgartner, 2001), RoadRunner (Crescenzi et al, 2001) and WebTable (WebTable, 2022)) perform IE on semi-structured online documents such as HTML pages, primarily using the layout structure of the documents. In contrast, the IE tasks are those which process free text in natural language, since this "unstructured" data is the primary source of human-generated information (Sarawagi, 2006).

Related Work on Named Entity Recognition

This research proposes a novel task of value identification in the frame of Value Proposition as a part of the existing techniques of Named Entity Recognition (NER), where the word or a phrase expressing value is treated as a Named Entity.

The main area of the information extraction technique described in this chapter is NER and the most relevant related works are summarized in this section.

(Ortona, 2016) concludes that three groups of studies on Named Entity Recognition could be determined at present, based on the information source which is leveraged for producing text annotation:

Language-driven approach presents traditions in Named Entity Recognition. It exploits knowledge of natural languages to obtain linguistic pattern normally forming the rules which applying to texts retrieve instances of entities of interests;

Knowledge-driven approach utilizes structured, i.e., ontological or unstructured, i.e., gazetted, background information locating an entity of interest in the text documents;

Combining annotators. Usually approaches that tackle the aggregation of multiple NER systems have exploited voting mechanism or variations of it (Kozareva, 2007).

As general, those techniques frequently do not pay attention to certain information regarding text semantics or contexts. An approach, namely Nerd (Nerd, 2022), proved that ontological background knowledge - this managed to be useful with creating annotations - could as well be helpful with aggregation of numerous annotators, logically specifying incoherent annotations.

Value Entity Recognition (VER)

Typically, named entities are a type of proper names while others are general entities within the domain. Thus, NER is processed based on orthographic features,

word-level features, and gazetteers (Ngo, 2021). To classify named entity tags and the disambiguation among named entities based on their features, data analytic techniques are used. However, these features have very little contribution to the Value Entity Recognition (VER) task because, as (Ngo, 2021) showed based on the Agricultural Entity Recognition (AGER), general entities are not proper names and they do not have orthographic features like named entities in the NER task. For entity recognition systems, it is a challenge.

This dissertation proposes an entity recognition process in a specific domain which is air transport, and the entity is a value as a part of value proposition. The key highlights of the VER are: (1) design rule-based annotation experiment for value entities; (2) create a semantic-based Value Entity Recognition (VER) task and a original effective approach for VER in air transport domain; (3) utilize deep learning (DNN) to train and identify value entities from texts of air transportation Twitter profiles of start-up companies. The semantic-based VER approach combines a NER-based approach and binary classification approach (if an entity is a value or is not a value). Figure 2 shows the place of Value Entity Recognition (VER) as an NER approach to Information Extraction Process (adopted from Alexander, 2013).

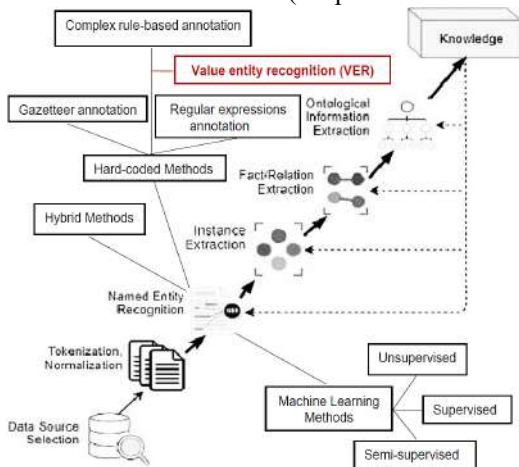


Figure 2. Value Entity Recognition (VER) as an NER approach to Information Extraction Process (based on Alexander, 2013)

Determining the methods of identifying values from texts, several alternatives were evaluated as reflected in Table 2.

Table 2. Alternative methods for value identification from texts

Method of Information Extraction	Reason for Accepting / Rejecting
• Multiclass (speed, cost, comfort), predetermined classes	• numerous classes and no opportunity to find new classes/values
• sentiment analysis	• whole text/sentence is evaluated; multiclass (values can belong to a mixture of sentiment classes)
• binary term classification	• allows to identify new values
• Named Entity Recognition (NER)	• allows to create a new entity for recognition

NER approach allows to create *a new entity type - value* with predefined semantic rules for annotators: they learn a basic theory of a value proposition and they are trained to spend the same time as an average online consumer spends on a webpage - 62 seconds (Contentsquare Homepage, 2020).

Concluding remarks

1. The non-data-driven traditional tactic suffers from several disadvantages, that limits effectiveness in decision-making: human inputs, biases and subjectivity, prone to errors by human decision-makers, typically be more expensive, time-consuming.
2. To support a business process of delivering new value proposition, this research offers data-driven approach to identification of values, increasing numbers of values, and automation of this process.
3. Value proposition of the company in vast majority of cases is stated in the natural language and contains values delivered by words.
4. Aviation English related works analysis identifies a gap of the research in the following areas: texts from webpages, official regulations, companies' policies, work reports, research papers in the domain of Air Transport to be utilized as databases for values identification and number increasing.
5. From the Natural Language Processing perspective, values in the texts of value propositions can be considered as named entities and be identified using the technique known as Named Entity Recognition.

This dissertation proposes an entity recognition process in a specific domain which is air transport, and the entity is a value as a part of value proposition. The key highlights of the VER are: (1) design rule-based annotation experiment for value entities; (2) develop a semantic-based Value Entity Recognition (VER) tasks and a new efficient approach for VER in air transport domain; (3) use deep learning (DNN) to train and identify value entities from texts of air transport Twitter profiles of start-up companies. The semantic-based VER approach combines a NER-based approach and binary classification approach.

12.3. Methodology and Research Design

In this Chapter, the methodology and research design are discussed. The presented research is novel from the perspective of a set of existing methods, applied to a new task: to identify and expand the values in air transportation and to show the possibilities of automating the process of determination and expansion. A novel "*One-Word principle*" has been applied for the analysis of data.

Research Approach and Methods

A research procedure is divided into *seven steps*:

- 4) 1st step is the theoretical analysis on values aimed at determining meta-characteristics of the process of value identification and expansion in the air transportation domain. It includes: Secondary data analysis as review of theoretical literature, previous research, policy documents; Framing as evaluation and conceptualization of approaches and notions

5) 2nd step determines ending conditions that should be met: values expressed by words are recognized and categorized, experts' agreement is achieved and a model created

6) 3rd step adopts two approaches employed in the research: ***empirical to conceptual*** and ***conceptual to empirical***.

Empirical to conceptual

4a) to identify subset of objects, a research base is determined as landing pages of start-up companies in the air transportation domain; values are annotated, dataset is created as a bottom-driven approach without prior conceptualization.

5a) to identify common characteristics and to group objects, semantic analysis of values is performed, aimed at potential categorization.

6a) to group characteristics into dimensions to create taxonomy, categories are examined and arranged hierarchically into subclasses and then classes.

Conceptual to empirical

4b) prior conceptualization of the characteristics and dimensions of objects: it is assumed that values in the texts can be identified and expanded automatically.

5b) to examine objects for these characteristics and dimensions, a neural network model was trained.

6b) a final dataset is received and accuracy is calculated.

7) confirm that final conditions are met.

To develop the research procedure, modern approaches to Information Extraction from texts were examined and the following conclusion was made: the metadata extracted by IE tasks for understanding the raw text is treated as the semantic annotation, and the text analysis executors -humans or programs - tackling IE tasks as semantic annotators.

Qualifying Value Entity Recognition as a task for semantic annotators, it was necessary to analyze the potential of different IE approaches for a novel task of value identification:

–*Named entity extraction*: recognize and classify atomic snippets in text as referring to particular semantic categories (e.g., Person and Company) (Cunningham, 2002). In some cases, it also involves a subtask named entity disambiguation, which resolves entity identifiers of the entity instance detected (Li et al, 2013);

–*Relation extraction*: extract relationships between entities (Reichartz et al, 2010);

–*Sentiment annotation*: identify viewpoints or emotions in the underlying text. The attitude or tone can be polarity opinions such as Positive/Negative or comparison opinions such as Better/Worse or simply Subjective/objective (Reichartz et al, 2010);

–*Topic annotation*: attach documents with topics or themes (Silva et al, 2001);

–*Co-reference annotation*: catch co-referring mentions of a given entity (Deemter and Kibble, 1995);

–*Keyword annotation*: extract relevant keywords or terminology to the content the text (Liu et al., 2009).

From the Natural Language Processing perspective, values in the texts of value propositions can be potentially considered as named entities and be identified using the technique known as Named Entity Recognition. After the thorough evaluation of the nature of values as parts of value proposition (see 2.5), the new type of entity introduction - Value Entity Recognition (VER) - was introduced.

The above-described *approaches* have led to the choice of ***methods used for data collection***:

- Semantic annotation combined with a binary classification approach (if an entity is a value or is not a value) for landing pages labelling; Twitter profiles of the start-ups labelling.
- Survey – resulting from self-formulated answers to the open question *What values do you expect from air transportation industry?*

The following ***data-analysis methods*** were applied for systemizing the obtained data: categorization and organization into taxonomies.

A corpus of words constitutes an uncontrolled vocabulary, e.g., uncategorized and not organized for data retrieval. To use a vocabulary in thesauri, ontologies, taxonomies, and other knowledge organization schemes, the initial uncontrolled vocabulary is organized into a *controlled vocabulary* - controlled vocabulary is normally utilized for describing contents by allocating terms to represent metadata associated with content objects, organizing content on websites and the alike (University of Texas, 2022).

There are four important principles of vocabulary control that guide their design and development. To create controlled vocabularies, four major principles exist that help designing and developing process: *to eliminate ambiguity, to control synonyms, to establish relationships of terms, to test and validate* (ANSI/NISO Z39.19-2005). It is possible to identify many forms of semantic links between terms in controlled vocabularies. Those involve associative relationships, hierarchical relationships, and equality connections that can be described according to the needs of a given implementation.

Procedure of the Experiment

To identify and expand values from texts of air-transportation start-up, there was conducted an experiment, consisted of **2 stages**. The 1st stage includes a survey and start-ups annotations (Figure 3) and the 2nd stage includes annotation of the Twitter profiles of start-ups from the 1st stage. The main technique used is annotation. Data annotation is a text-mining technique through which data is labelled to make objects (in this case - values expressed by words) recognizable by machines.

Research base

Experiment is conducted on the start-ups texts as a data source of values.

Experiment, Stage 1. 796 start-ups' landing pages of air transportation domain, source: www.crunchbase.com.

Experiment, Stage 2. 796 twitter profiles of the start-ups from Stage 1 of air transportation domain, source: www.crunchbase.com

1st stage: Education, Pre-survey, Training, Annotation, Post-survey (Figure 3).



Figure 3. 1st stage of procedure

Annotators were educated about the value proposition concept, conducted a survey with open formulation of values in AT, trained to label values, labelled values and conducted a post-survey to find a difference with pre-survey for future research.

An online interface was created (Figure 4) and a database of 796 start-ups landing pages from air transportation domain was downloaded.

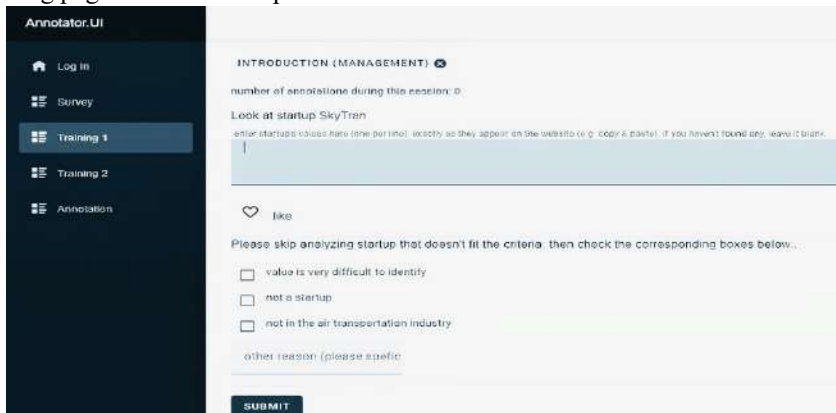


Figure 4. Online interface for Stage 1 - landing pages annotation

96 annotators from different background with no professional experience of value identification and English proficiency on the level of B2+ according to Common European Framework of Reference (CEFR) were tasked to annotate texts for values.

The aim of a preliminary survey is to understand the level of expertise for non-trained participants and to determine value proposition volume and structure before the education process. The students were told about the basic concept of value proposition. After that they were asked to list as many values as they can, that are provided by the Air Transportation start-ups and companies.

After the participants were told the start-up concept, they learned existing Value Proposition major theories. After that they were offered two trainings. Training 1 shows five relatively simple examples of how to determine Value Proposition on start-ups landing pages. Training 2 presents five landing pages where the Value Proposition cannot be easily identified due to its vague expression or non-standard ways to promote their start-up, e.g., video.

The participants have options to identify a landing page as not a start-up, can state that the value is very difficult to identify, and can name the page as not from the air transportation industry. Also, they are asked to click a Like button if they think this start-up clearly delivers the Value Proposition.

As the result of the experiment, rich semantic metadata has been created. Two databases were collected: corpus of values from open-formulated survey and annotated values from texts of AT start-ups. The quality of annotation was re-assessed by industry experts.

Demography of Annotators

1st Stage In order to collect generalizable observations on air transportation value proposition, for the 1st Stage there were recruited 96 students from the transport university with a connected, but still diverse range of majors and/or professional roles in a transportation field: Aviation Management, IT, Logistics. All participants were between 20-42 years of age.

2nd Stage Second stage of the experiment is consisted of annotation of values from Twitter profiles. Twitter profiles present 796 start-up companies, which landing pages were annotated in the 1st stage of the experiment. Two experience annotators with the deep knowledge of value proposition concept were employed. A separate online interface for annotation with profiles on the screen was created as shown on Figure 5.

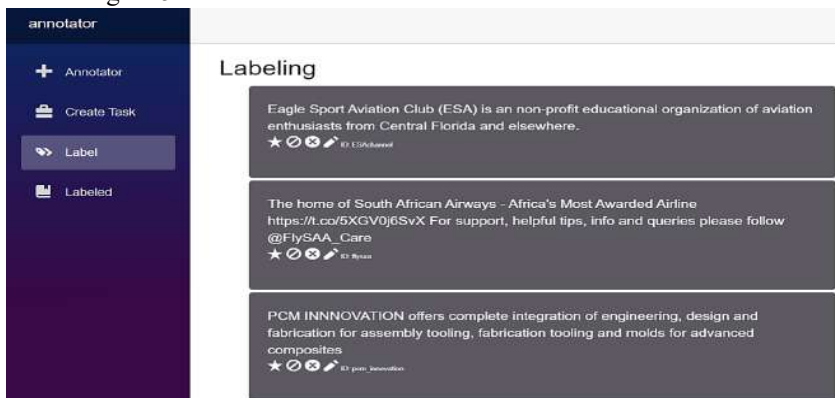


Figure 5. Online interface for Stage 2 - Twitter profiles annotation

As a result, a corpus of values from Twitter profiles of air transportation start-ups was collected.

Taxonomy Development Approach

Two taxonomies have been developed based on the received data: Air Transport Value Taxonomy and Semantic Taxonomy of Values in Aviation.

Air Transport Value Taxonomy was developed on the Pre-Survey data. Semantic Taxonomy of Values in Aviation (STVA) was developed on the start-ups landing pages annotation data.

Semantic Taxonomy of Values in Aviation also aligned with International Organization for Standardization terms and was validated through interaction with

air transportation experts: one IT director of the international airline and two commercial directors of airports.

Based on the approach and methods used two taxonomies of values in air transportation were developed. Below, the author presents the specific principle of data analysis applied in the process of Semantic taxonomy development and NLP model implementing – *One-Word principle*. As a part of a current research aiming at automated value recognition, the author assumes that the majority of values could be captured by one word.

The validity of the assumption is investigated via experiments in the domain of aviation: the analysis of the Tweets annotation shows that more than 50% of all the values identified can be captured by one word: 73% of values were annotated as one-word value.

12.4. Key Findings and Research Results Implementation

The obtained findings are as follows.

- 1) *Conceptual analysis and identification of concepts* (building a conceptual framework) were conducted:
 - Changes in the focus of research on Aviation English for domains in 2002-2016 and 2017-2021 were specified.
 - Air transportation value shift through history was identified.
 - Etymological genesis of the term *value proposition* was developed.
 - Key factors influencing value proposition in aviation were identified.
- 2) *The corpus of values in air transportation consisting of 3 parts* (survey, startups annotation and Twitter profiles annotation) was constructed and a Neural Network model was trained
- 3) *One-word approach* for capturing values in the texts was conceptualized
- 4) *Two classifications of values* in air transportation domain were constructed:
 - Air Transport Value Taxonomy, customer perspective
 - Semantic Taxonomy of Values in Aviation, industry perspective
- 5) Methodology of automated identification of values from texts is developed

Conceptual Analysis and Identification of Concepts

To build a conceptual framework of the *value proposition* notion and *values* as part of it, to identify the role of values in Air Transportation strategic management, and to review the literature on the research of English language usage in Air Transportation, a series of studies was performed with the results described in this chapter.

Results of conceptual analysis and identification of concepts are presented in the theoretical part of this thesis *Values as Part of Value Proposition in Air Transportation Industry*:

- Changes in the focus of research on Aviation English for domains in 2002-2016 and 2017-2021 were specified
- Air transportation value shift through history was identified
- Etymological genesis of the term *value proposition* was developed
- Key factors influencing value proposition in aviation were identified.

The Corpus of Values in Air Transportation Corpus

A corpus of values in air transportation domain is created consisting of three parts:

1. Survey based and resulting from self-formulated answers to the open question What values do you expect from air transportation industry? with 241 values.
2. Annotation based and resulting from start-ups' landing pages labelling with 588 values.
3. Annotation based and resulting from Twitter profiles of the start-ups labelling with 763 values.

Implementation with neural network

Based on Experiment, Stage 2.

Semi-supervised method of value extraction was applied with the following steps:

1. Set of documents (564 Twitter profiles of aviation start-ups) was used as a source base for value recognition
2. Labelling (annotation of values) by experts assumed words containing values as part of value proposition to be tagged as value entities. Combined method of semantic and binary annotation was utilized.
3. Dataset of 763 values was finally obtained as the result of annotation.
4. Pre-processing (filtering, normalization: Python) included One-word principle: one value is expressed by one word.
5. Simplifying task as binary classification: binary classification (value / not value) instead each value being a class, allows to pick up new values.
8. A neural network model from spaCy library is used as a binary classifier.
9. Training and evaluation: Figure 6 presents the process of training spaCy model.

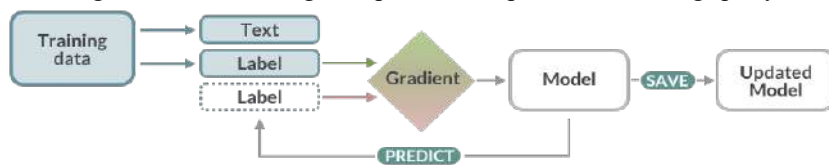


Figure 6. Process of training the model

The word corpus used is "spacy.Corpus.v1" for English language.

Model architecture is Spacy.TransitionBasedParser.v2. Transition-based parsing is an approach to structured prediction where the task of predicting the structure is mapped to a series of state transitions.

Architecture:

Hidden layer size = 64 neurons. Parser consists of 3 subnetworks:

1. tok2vec layer (transforms words (tokens) into vectors)
2. "lower" layer (the "meat" of neural network, converts vectors to internal representation)
3. "upper" layer (feed-forward layer that calculates scores from the previous layers' vectors)

The scores then are used to predict the label for the word.

Figure 7 shows the training progress, where *ents_p*, *ents_r*, *ents_f* are the precision, recall and f-score for the NER task.

```

===== Training pipeline =====
[i] Pipeline: ['tok2vec', 'ner']
[i] Initial learn rate: 0.001
E   #           LOSS TOK2VEC   LOSS NER   ENTS_F   ENTS_P   ENTS_R   SCORE
-----
0   0           0.00         75.80     55.76   49.89   63.20   0.56
0   10          0.99         710.78    71.61   61.04   86.59   0.72
0   20          3.28         451.60    72.88   62.92   86.59   0.73
0   30          3.09         400.33    76.67   68.96   86.32   0.77
0   40          5.55         289.59    81.32   78.57   84.27   0.81
0   50          9.35         267.72    82.42   80.03   84.95   0.82
0   60         11.34         254.40    82.69   79.70   85.91   0.83
0   70         10.72         225.34    83.03   79.97   86.32   0.83
0   80          9.20         236.51    81.70   80.67   82.76   0.82
0   90          6.87         244.33    84.25   81.30   87.41   0.84
0  100          7.37         247.18    84.38   81.08   87.96   0.84

[+] Saved pipeline to output directory
spacy_output\model1-last

[2021-11-10 16:17:39,067] [INFO] Set up nlp object from config
[2021-11-10 16:17:39,076] [INFO] Pipeline: ['tok2vec', 'ner']
[2021-11-10 16:17:39,079] [INFO] Created vocabulary
[2021-11-10 16:17:39,081] [INFO] Finished initializing nlp object
[2021-11-10 16:17:39,753] [INFO] Initialized pipeline components: ['tok2vec', 'ner']

```

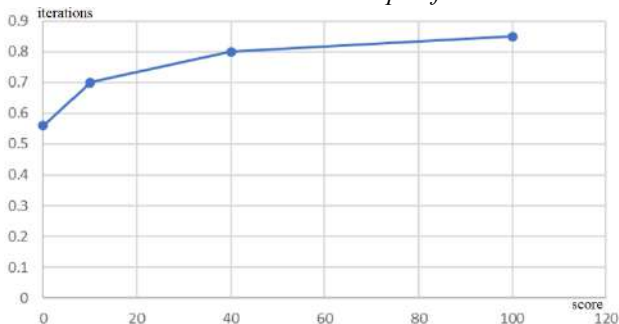
Figure 7. Pipeline

Start from Score 0.56, after 10 training iterations the score raised up to 0.72, after 60 training iterations the score raised up to 0.83.

The experiment was stopped at 100 iterations with the score at 0.84 because of the following reason:

1. improvement flattens out (becomes incremental) (see Table 3)
2. to avoid overfitting

Table 3. Graph of iterations / score correlation



Metrics are as follows: precision - 0.85; recall - 0.89; f1-score = 0.87

The conducted experiment allowed collecting an extended database of values. To systemize the data, two taxonomies were created:

- *Air Transport Value Taxonomy*, customer perspective, survey-based
- *Semantic Taxonomy of Values in Aviation*, industry perspective, annotation based

Air Transport Value Taxonomy: description

A survey, conducted using an open-ended question *What values do you expect from air transportation industry*, allowed to obtain a dataset of self-formulated 241 values. To categorize them, a taxonomy was constructed. The adopted survey methodology with open-ended questions and self-formulated answers allows to attribute this taxonomy to reflecting the consumer's perspective.

Semantic Taxonomy of Values in Aviation: description

An annotation of texts in landing pages of start-up companies in air transportation domain allowed to collect a dataset of 588 values. Using a frequency threshold equal to 12, 196 values were captured for the following analysis. To categorize them, a taxonomy was constructed (Figure 8). The methodology of annotating texts created by companies allows to attribute this taxonomy to reflecting the industry's perspective.

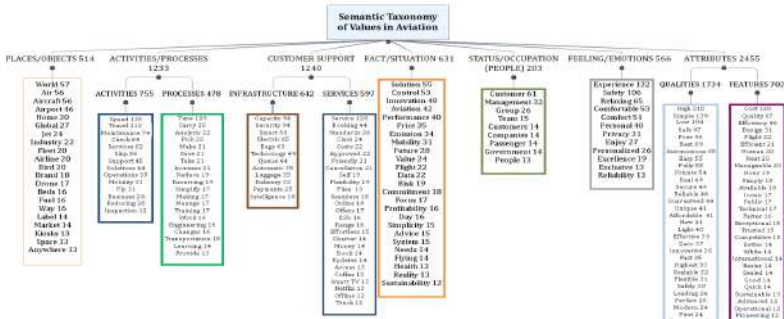


Figure 8. Semantic Taxonomy of Values in Aviation

Taxonomy groups of all levels are not tied to specific parts of speech and are united by the same semantic concept.

Categories are not collapsed by excluding partial synonyms as to preserve semantic difference might be important for marketers and customer service. Taxonomic classes and subclasses were developed from the 196 values describing activities, procedures, objects, emotional experience and considering the context in which they could be possibly associated with in air transportation.

By analyzing the terms and the context, seven taxonomic classes and six subclasses have been developed. Each of the 196 lowest level terms fits into at least one taxonomic group. However, many of them belong into more than one taxonomic group and was assigned to a specific group under the five experts' agreement: two of the experts have degree in Linguistics and 20+ years of linguistic research, one expert possesses a PhD and 20+ years in Natural Language Processing, one expert is a top manager of the regional airport holding PhD, and one expert is Information

Technologies Director of International Airline with 20+ years of experience in air transportation. Categories are not mutually exclusive, but this taxonomy is not aimed to provide overlapping semantic clustering.

Every group and each term present a number of people that name this value. The most frequent terms (stop words excluded) are: high (215), cost (160), and simple (53). The weight of each group is a sum of its components. The biggest group is Attributes with 2455 and the smallest one is Status / Occupation (People) with 203.

Proposed Methodology of Automated Identification of Values from Texts

Figure 9 presents the final result of the research and the achievement of the research aim – the Methodology of Automated Identification of Values from Texts.

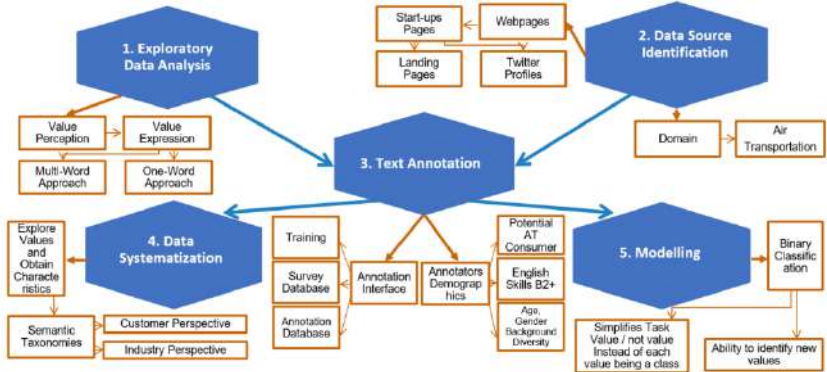


Figure 9. Proposed methodology for automatic identification of values in the domain of air transportation (author’s made)

The proposed methodology is limited by one domain – Air Transportation as Air Transportation domain was accepted and tested in the case study in this research and the developed Natural Language Processing model can be implemented to the texts of Air Transportation.

Nevertheless, the methodology is applicable to any domain and if follow all the steps consistently since the methods of semantic annotation and data categorization have no domain restrictions and start-up companies as data source cover most industries.

12.5. Research Results Discussion

The analysis presented in this dissertation supports the initial thesis that values as a part of value proposition can be automatically identified and number of values be increased. Additionally, this research seeks to contribute to the existing practice and literature both empirically and theoretically.

Theoretical and empirical contribution

These analyses increase the theoretical understanding of the value notion, value comprehension by the customers, and theoretical framework of the methods of value automated identification and expansion.

This research is the first to conduct a task of automated value identification and expansion. It provides the theoretical framework of automated value identification and expansion that can be applied to any domain.

This study is the first considering values as a part of value proposition as the primary field of investigation. Additionally, this study uses a novel approach for taking texts of start-up companies as a research base for identifying and expanding values. This broadens the existing techniques on developing value proposition of a company. Moreover, the obtained data show that values can be significantly extended and categorized to support the decision-making process. Through this, it was identified that deep learning tools can be utilized to automate the process of value identification and expansion.

Practical Importance and Industry Reviews

Three air transportation industry experts of a high level of expertise - two commercial directors of international airports and one IT head manager of the major airline - were asked to compare the developed Value Proposition taxonomy with IATA Airline taxonomy. They also shared their ideas about the way taxonomies can be useful for industry and COVID-19 surprisingly little significance from respondents in April, 2020. Below there is a summary of the experts' review:

Air Transportation industry experts' reviews:

International Hub Airport Commercial Director: "The unexpected finds and formulations of values that were not previously in sight are of great interest for marketers. Good perspective can be seen in using aviation values taxonomies in Project Management"

Airline IT Head Manager: "These taxonomies are extremely useful for the following purpose: to understand the essence and systematization of services / products. Our company has used a unique Semantic Taxonomy of Values in Aviation while developing our new online check-up and shopping system"

Commercial Director of Reginal International Airport: "Automation and systematization of values in air transportation allows for faster decision making on strategic questions as it shows new values and value chain transformation; taxonomies enable to formulate Key Performance Indicators by directions, to distribute areas of responsibility between all the links in the value chain."

Conclusions and Recommendations

Conclusions

This study has been designed to help companies use the potential of data processing to create new data-driven decision support systems in the process of developing a value proposition. ***The research aim to develop a methodology for automatic identification and expansion of values in the domain of air transport*** has been achieved, and methodology has been elaborated.

The main research question Can values be identified, and number of values be increased automatically? has been answered affirmatively.

Five chapters of this Thesis have developed new methodology of value as part of value proposition identification and extension as well as a model and show

how these allows to yield concrete results in the automation of the process of value identification and extension.

To prove the usability of the value automatic identification and expansion, modern perspectives on the deployment of data-driven business processes have been discussed. The benefits of utilizing data analytics adopting Natural Language Processing approach and text mining techniques to enhance business performance have been demonstrated: time- and effort-saving. Reviews from leading industrial analysts have been provided, including insights on the advantages of a digital strategy for business processes and implementation of data-driven decision-making support systems: automation and systematization of values in air transportation allows for faster decision making on strategic questions.

The importance of novel and unique values is illustrated by the following conclusions:

- in the face of growing competition and air transportation liberalization, the role of a unique and novel value proposition in aviation is increasing and fast becoming a strategic priority.
- the traditional values in air transportation are price, quality and speed. In the recent decades, commercialization of air transportation initiates competition and, subsequently, new values search.
- value creation is one of the main company goals, and after the recent liberalization air transportation industry faces the unmet need of identifying new value proposition. Automatization of this process would be beneficial.
- to automate the process of value recognition, start-up companies present a decent data source for analysing value proposition since their goal is to grow and scale rapidly, information on their landing pages is open and actual.

The final methodology of automated value identification includes the following steps: Exploratory data analysis, data source identification, text annotation, data systematization, and modelling.

The author views this work as a proof of concept, showing that there is a place for reasoning about automation for value identification and expansion and developing a model. The most immediate application of this tool is envisaged in air transportation companies developing new value proposition in the current challenging wellbeing and political environment.

Future research can be devoted to the following:

1. including different domains into the corpus of values
2. applying the resulted methodology through the certain periods of time aimed at obtaining time series of values to widen value metrics and to follow the value shift
3. to identify values from official regulations could provide some insights on official policies different viewpoints comparing to those of the industry.

Reproducibility of the research can be proven by the following the steps of methodology and acquiring the high degree of reliability through evaluating the results and accuracy.

Recommendations

Based on the results obtained, discussions with experts and conclusions, the author of this study has formulated the following recommendations:

To decision-makers in the Air Transport domain: automatically identifying values as part of a value proposition is beneficial for significant reduction of the time and effort to comprehend existing values in the market and to discover new demand trends

To marketers in the Air Transport domain: the wording of respondents' answers to open-ended questions about expected values from aviation supports to understand what and how the consumer expects from the air transport industry.

To industry analysts: automatic valuation enables lower-cost analytical insights into the market in the context of the value proposition, strategic planning, and new products trends.

Autora publikāciju bibliogrāfiskais saraksts/ Publications with author's participation

1. Zervina O., 2022. *Value Expansion and Sense Making*. BHMK-D-22-00023R1. Behaviormetrika, Japan ISSN: 1349-6964, DOI: 10.1007/s41237-022-00179-7 (Scopus)
2. Zervina, O. and Stukalina, Y., 2022. *Determining Value Proposition In The Context Of Aviation Industry Strategic Management*. Scientific Programme Committee, p.270. The 14th International Scientific Conference “New Challenges in Economic and Business Development - 2022: Responsible Growth”: Riga, Latvia, May 13, 2022. Proceedings. Riga: University of Latvia, 2022, 275 p (in process Web of Science, EBSCO)
3. Zervina, O., 2021. *The environment of extremist textual content threatening transportation systems*. In International Conference TRANSBALTICA: Transportation Science and Technology (pp. 541-551). Springer, Cham. DOI: 10.1007/978-3-030-38666-5_57 (WoS)
4. Zervina, O., Stukalina, Y., Pavlyuk, D. and Rubens, N., 2021, October. *One-Word Approach in Text-Mining for Value Identification*. In International Conference on Reliability and Statistics in Transportation and Communication (pp. 135-143). Springer, Cham. DOI: 10.1007/978-3-030-96196-1_13 (Scopus)
5. Neil Rubens, Olga Zervina, Yulia Stukalina, 2020. *Air Transport Value Taxonomy*. 15th International Forum on Knowledge Asset Dynamics (IFKAD 2020) - 09-11 Sep 2020, University of Basilicata, Italy, pp. 178-194. ISBN: 978-88-96687-13-0 (Web of Science)
6. Zervina, O., Stukalina, Y., Pavlyuk, D. and Rubens, N., 2020, October. *Value creation in air transportation: beyond price, quality, and speed*. In International Conference on Reliability and Statistics in Transportation and Communication (pp. 119-129). Springer, Cham. DOI: 10.1007/978-3-030-68476-1_11 (Scopus)

7. Zervina O., Rubens N., 2020. *Air Transport Value Proposition*. COVID-19 Affect. XXIII Annual International Professional Conference, Sting Academy, 2020, BRNO, Czechia ISSN 1805-6873 EBSCO, (ERIH PLUS)
8. Zervina O., 2020. *A Linguistic Analysis of Startups in The Context of the Air Transport Industry Management*. 14th Baltic DB&IS, June 16-19, 2020, Tallinn, Estonia. <http://ceur-ws.org/Vol-2620/> pp.57-64
9. Zervina, O., 2019. *The Analysis of Threat Defining Language in the EU Official Transport Security and Safety Documents*. In International Conference on Reliability and Statistics in Transportation and Communication (pp. 329-338). Springer, Cham. DOI: 10.1007/978-3-030-44610-9_33 (Scopus)
10. Zervina O. and Stukalina Y., 2022. *Factors Influencing Value Proposition in the Aviation Industry in the Context of Customer-Centric Digital Economy*. RelStat-2022, TSI, Riga, Latvia (accepted for conference presenting and publishing)
11. Zervina O. and Stukalina Y., 2022. *Using AI for scaling up a business in the age of digital transformation*. RelStat-2022, TSI, Riga, Latvia (accepted for conference and publishing)
12. Zervina O., Stukalina Y., and Pavlyuk D., 2022. *Value Entity Recognition Task in the Air Transportation on the Base of E-Texts Analysis*. RelStat-2022, TSI, Riga, Latvia (accepted for conference presenting and publishing)

Literatūras saraksts/References

1. Alexander F.: "Information extraction - lecture 3: Rule-based named entity recognition", In: Lecture slides, published at https://www.cis.unimuenchen.de/fraser/information_extraction_2015_lecture/, accessed on 10/11/2020, pp. 512-517 (2013).
2. Almquist, E., Cleghorn, J. and Sherer, L., 2018. The B2B elements of value. Harvard Business Review, 96(3), p.18.
3. Almquist, E., Senior, J., Bloch, N. 2016, "The Elements of Value", Harvard Business Review, pp. 46-53
4. Äyväri, A. and Jyrämä, A., 2017. Rethinking value proposition tools for living labs. Journal of Service Theory and Practice. Vol. 27, No. 5, p. 1024-1039
5. Bain&Company, 2011. "Management tools: strategic planning", Bain&Co. Publications, [Online], available at: www.bain.com/publications/articles/management-tools-2011-strategicplanning.aspx.
6. Barnes, C., Blake, H. and Pinder, D., 2009. Creating and delivering your value proposition: Managing customer experience for profit. Kogan Page Publishers.

7. Baumgartner, R., Flesca, S. and Gottlob, G., 2001. Visual web information extraction with lixto.. In Proceedings of the 27th International Conference on Very Large Data Bases, VLDB '01, pages 119-128, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.
8. Chen, L., 2014. Integrating and Querying Semantic Annotations, doctor dissertation. Oxford University press.
9. Contentsquare Homepage, 2020. Digital Experience Benchmark, <https://go.contentsquare.com/en/digital-experience-benchmark>, last accessed 2021/07/07.
10. Crescenzi, V., Mecca, G. and Merialdo, P., 2001, September. Roadrunner: Towards automatic data extraction from large web sites. In VLDB (Vol. 1, pp. 109-118).
11. Deemter and Kibble, R., 1995. On coreferring: coreference in muc and related annotation schemes. *Computational Linguistics*, 26:629-637, 1995.
12. Ebel, Philipp; Bretschneider, Ulrich; Leimeister, Jan Marco (2016): Leveraging virtual business model innovation. A framework for designing business model development tools. In *Information Systems Journal* 26 (5), pp. 519-550. DOI: 10.1111/isj.12103.
13. European Commission, 2016. Europe's Next Leaders: the Start-up and Scale-up Initiative. https://ec.europa.eu/growth/content/europes-next-leaders-start-and-scale-initiative_enGiardino et al, 2014
14. Hein, A., Weking, J., Schreieck, M., Wiesche, M., Böhm, M., and Krcmar, H., 2019. "Value co-creation practices in business-to-business platform ecosystems," *Electronic Markets* (29:3), pp. 503-518. (PDF) Challenges in the Emergence of Data Ecosystems. Available from: https://www.researchgate.net/publication/341930759_Challenges_in_the_Emergence_of_Data_Ecosystems [accessed Apr 2022].
15. IATA Passenger Surveys, 1991-2019. IATA Global Passenger Highlights, <https://www.iata.org/contentassets/952a287130554b4880563edca1c8944f/iata-2019-gps-highlights.pdf>
16. IATA, 2018. Future of the Airline Industry 2035. International Air Transport Association. [Online] Available from: <https://www.iata.org/contentassets/690df4ddf39b47b5a075bb5dff30e1d8/iata-future-airline-industry-pdf.pdf> [Accessed 10.09.2021]
17. Kaplan, R. S., and Norton, D. P., 2008. The execution premium: Linking strategy to operations for competitive advantage. Harvard Business Press.
18. Kozareva Z., Fernandez O., Montoyo A., Munoz A., Suarez A., and Gomez J., 2007. Combining data-driven systems for improving named entity recognition. *Data & Knowledge Engineering*, 61(3):449-466, 2007.
19. Lanning and Michaels, 1988. Delivering value to customers. Harvard Business Review, President and Fellows of Harvard College.
20. Li Yang, Chi Wang, Fangqiu Han, Jiawei Han, Dan Roth, and Xifeng Yan, 2013. Mining evidences for named entity disambiguation. In Proceedings of the 19th ACM SIGKDD International Conference on Knowledge

- Discovery and Data Mining, KDD '13, pages 1070-1078, New York, NY, USA, 2013. ACM.
21. Liu, F., Pennell, D., Liu, F. and Liu, Y., 2009, June. Unsupervised approaches for automatic keyword extraction using meeting transcripts. In Proceedings of human language technologies: The 2009 annual conference of the North American chapter of the association for computational linguistics (pp. 620-628).
 22. Maslow, A., 1943. A theory of human motivation. *Psychological Review*, 50(4), 370-396. Doi:10.1037/h0054346, (1943).
 23. Nerd, 2022. Named Entity Recognition and Disambiguation. <http://nerd.eurecom.fr/documentation>
 24. Ngo, Q. H., Kechadi, T., & Le-Khac, N. A., 2021. Domain Specific Entity Recognition With Semantic-Based Deep Learning Approach. *IEEE Access*, 9, 152892-152902.
 25. Oliveira, M. I. S., and Lóscio, B. F. 2018. "What is a data ecosystem?" in Proceedings of the 19th Annual International Conference on Digital Government Research, A. Zuiderwijk and C. C. Hinnant (eds.), Delft, Netherlands. May 30-June 1, New York, USA: ACM.
 26. Ortona, S., 2016. Easing information extraction on the web through automated rules discovery. Doctoral dissertation, University of Oxford.
 27. Osterwalder, A. and Pigneur, Y., 2010. *Business Model Generation: a handbook for visionaries, game changers, and challengers*: John Winley & Sons.
 28. Prasada, P., Febrian, E., Komaladewi, R. and Zusnita, W., 2021. Competitiveness of SMEs: The support of value creation and market orientation. *Management Science Letters*, 11(2), pp.645-656.
 29. Reichartz, F., Korte, H. and Paass, G., 2010, July. Semantic relation extraction with kernels over typed dependency trees. In Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 773-782). Rintamäki et al, 2007):
 30. Sarawagi S., 2006. Scalable information extraction and integration. In Proceedings of the 13th International Conference on Management of Data, COMAD '06, page 249. Tata McGraw-Hill Publishing Company Limited, 2006.
 31. Silva, J., Mexia, J., Coelho, A. and Lopes, G., 2001, November. Document clustering and cluster topic extraction in multilingual corpora. In Proceedings 2001 IEEE International Conference on Data Mining (pp. 513-520). IEEE.
 32. Su-C Li, 2007. The Role of Value Proposition and Value Co-Production in New Internet Startups: How New Venture e-Businesses Achieve Competitive Advantage," PICMET '07 - 2007 Portland International Conference on Management of Engineering & Technology, Portland, OR, 2007, pp. 1126-1132, doi: 10.1109/PICMET.2007.4349435. University of Texas, 2022

33. WebTable, 2022. <http://windowsutilities.com/webtable.html>, checked 2022/07/07.
34. Wildevuur, S.E., van Dijk, D., Hammer-Jakobsen, T. and Äyväri, A., 2015, July. Designing value to enable chronic disease management through information and communication technology: the People Value Canvas. In Third European Conference on Design4Health, Sheffield, UK.
35. Yoo, Y., Henfridsson, O., and Lyytinen, K., 2010. "Research Commentary —The New Organizing Logic of Digital Innovation: An Agenda for Information Systems Research," *Information Systems Research* (21:4), pp. 724-735