

LATVIJAS UNIVERSITĀTE
DATORIKAS FAKULTĀTE

MAIJA KĀLE

**Jauna konceptuāla ietvara izveide labākai izpratnei par pārtikas
patērētāju: starpdisciplināra lielo datu pieeja**

Promocijas darba kopsavilkums

Zinātnes nozare: datorzinātne

Apakšnozare: citas datorzinātnes un informātikas apakšnozares: uztveres un kognitīvās sistēmas

Zinātniskais vadītājs:
Dr. phil., profesors **Jurģis Šķilters**

Rīga, 2022

Promocijas darbs izstrādāts Latvijas Universitātes Datorikas fakultātē, Uztveres un kognitīvo sistēmu laboratorijā laika posmā no 2018. gada līdz 2022. gadam.

Atsauce: “LU doktorantūras kapacitātes stiprināšana jaunā doktorantūras modeļa ietvarā” nr. 8.2.2.0/20/I/006.

Promocijas darba kopsavilkums sastāv no terminu skaidrojuma, ievada, četrām nodaļām, nobeiguma un literatūras saraksta.

Darba forma: publikāciju kopa datorzinātnes nozarē, uztveres un kognitīvo sistēmu apakšnozarē.

Darba zinātniskais vadītājs *Dr. phil.*, profesors Jūrgis Šķilters.

Darba recenzenti:

1. Prof. Kārlis Čerāns (*Dr. sc. comp.*), Latvijas Universitāte
2. Prof. Ramešs Džeins, (*Dr. sc. comp.*) /Ramesh Jain, Professor of Computer Science/, Kalifornijas universitāte, ASV
3. Pētniecības vadītājs Mārcis Pinnis (*Dr. sc. comp.*), SIA Tilde, Latvija

Promocijas darba aizstāvēšana notiks 2023. gada 14. aprīlī plkst. 14.00 Latvijas Universitātes Datorikas fakultātē datorzinātnes un informātikas nozares promocijas padomes atklātā sēdē.

Ar promocijas darbu un tā kopsavilkumu var iepazīties Latvijas Universitātes Bibliotēkā Rīgā, Kalpaka bulvārī 4.

Latvijas Universitātes Datorikas fakultātes datorzinātnes promocijas

padomes priekšsēdētāja vietnieks _____ /prof. Jānis Bičevskis/ (*paraksts*)

padomes sekretāre _____ /Ruta Ikaunieca/ (*paraksts*)

Latviešu valodas korektore: Nadīna Elekse

No latviešu uz angļu valodu tulkoja: Māra Galeja

© Latvijas Universitāte, 2023

© Maija Kāle, 2023

ANOTĀCIJA

Autores motivācija šī promocijas darba ietvaros ir veidot jaunu pielietojumu lielajiem datiem par pārtiku. Datu daudzums par pārtiku pieaug, taču pārtikas starpdisciplināritātes dēļ daudzas nozīmīgas tēmas ir pētītas nozaru – tātad pārāk šauru disciplīnu – ietvaros. Autore uzskata, ka ar datorikas metožu palīdzību, analizējot lielos datus par pārtiku un ņemot vērā starpdisciplināru pieeju pētniecības jautājumu izstrādē, ir iespējams novērst šos trūkumus. Par to, ka tieši datorzinātnēm ir būtiska loma pārtikas patērētāju datu analīzē, liecina jaunas zinātnes nozares – pārtikas datorikas – attīstība, kuras ietvarā skatāms arī šis promocijas darbs. Ar pārtiku saistītu datu vākšanas un analīzes jomā ir vēl daudz neatrisinātu uzdevumu, tāpēc jauni veidi, kā izmantot sociālos medijus un citus liela mēroga datus, ir nozīmīgs jauninājums. Autore ir izstrādājusi vairākus veidus, kā izmantot liela mēroga datus, apvienojot kognitīvo un datorzinātņu metodiskos ietvarus, balstoties uz dabiskās valodas apstrādi un ietverot sekojošo metodoloģiju pielietojumu: a) sentimenta analīzi par afektīvām reakcijām uz pārtiku un multisensoro pieredzi saistībā ar pārtikas patēriņu, b) dažādu valstu virtuvju salīdzinošo analīzi, izmantojot tēmu modelēšanas metodoloģiju un c) bigrammu analīzi, lai izsekotu veidus, kā pārtikas patērētāji runā par veselīgu pārtiku. Promocijas darba rezultāti apkopoti jaunā konceptuālā ietvarā, ilustrējot tā pielietojumu praksē, kur ar dažādu datorzinātņu metodoloģiju palīdzību veikti vairāki pētījumi, kas apkopoti publikāciju sarakstā. Apvienojot sociālo, kognitīvo un datorzinātņu teoriju un metodoloģiju, jaunais konceptuālais ietvars nodrošina metodes un virzienu labākai izpētei pārtikas patērētāju pētījumos. Izmantojot datorzinātņu metodes lielu datu analīzei, un tā uzlabojot izpratni par pārtikas patēriņu, autores mērķis ir palielināt sabiedrības veselības politikas efektivitāti, tā nodrošinot labāku sabiedrības veselību un augstāku dzīves kvalitāti.

ANNOTATION

The author's motivation for this Doctoral Thesis is to create a new application of big data to food. While the amount of food data increases, many important topics have been studied within separate fields or too narrow disciplines, due to the interdisciplinary nature of this type of data. The author is of the opinion that these drawbacks can be eliminated by analysing food big data using computing methods and by taking interdisciplinary approach to develop research questions. The fact that computer science has a significant role in the analysis of food consumption data is evidenced by the development of a new field of science— food computing, which is the framework of this Thesis. Since there are still many unresolved tasks in the collection and analysis of food-related data, the use of social media and other type of big data is an important introduction. The author, by combining the methodological frameworks of cognitive and computer sciences, has developed several methods for using big data, based on the natural language processing and by applying the following methodologies: a) sentiment analysis of affective reactions about food and multi-sensory eating experience; b) comparative analysis of different national cuisines, by applying the topic modelling methodology; and c) bigram analysis to trace ways how food consumers talk about healthy food. Results of the Doctoral Thesis have been summarised in a new conceptual framework by illustrating its application in practice; in the course of application, several researches, which are specified in the list of publications, have been conducted using various computer science methodologies. As a combination of theories and methodologies of social, cognitive, and computer sciences, the new conceptual framework offers methods and a direction for studying food consumers. By using computer science methods in the analysis of big data and thus improving the understanding of food consumption, the author hopes to improve the efficiency of public health policies resulting in better public health and higher quality of life.

SATURS

SATURS	4
Terminu skaidrojums un definīcijas	5
IEVADS	6
Pētniecības joma un tās novitāte	7
Pētījuma motivācija, uzdevumi un mērķi	9
Pētniecības jautājums	9
Pētniecības metodes	10
Rezultātu aprobēšana	11
Publikāciju saraksts	11
Referāti konferencēs	11
Galvenie rezultāti	12
Praktiskais pielietojums	13
1. TEORĒTISKAIS PAMATOJUMS	15
1.1 Pārtikas uztveres un multisensoro pieredžu pētījumi	15
1.2. Pārtikas datorika un dabiskās valodas dati	16
2. PĒTĪJUMA PLĀNS UN METODES	18
2.1. Literatūras apskats	18
2.2. Dabiskās valodas datu analīze: datorlingvistika	20
3. KONCEPTUĀLAIS IETVARŠ	23
3.1. Izvēlētais tēmas konteksta zināšanu ieguve	23
3.2. Biežuma noteikšana	26
3.3. Emocionāli afektīvo reakciju noteikšana un laika skala	28
3.4. Papildu konteksta mainīgo pievienošana	30
3.4.1. Laikapstākļu dati	30
3.4.2. Veselīguma uztveres noteikšana ar datorlingvistikas metodēm	31
4. IEROBEŽOJUMI	33
NOSLĒGUMS	34
ATSAUCES	36
Interneta resursi	36
Literatūras saraksts	37
Pateicības	42

Terminu skaidrojums un definīcijas

Afektīvas reakcijas	Afektīvas reakcijas ir noskaņojums un emocijas, kas saistītas ar ēdienu. Šajā promocijas darbā afektīvas reakcijas attiecībā uz ēdienu tiek uzskatītas par svarīgiem virzītājspēkiem pārtikas izvēlē.
Pārtikas datorika	Pārtikas datorika ir starpdisciplināra joma, kas ietver ar pārtiku saistītus pētījumus, ko veic datorzinātnē (Min et al., 2019a). Lai izprastu dažādu ar pārtiku saistītu jautājumu pamatojumu, pārtikas datorika ir izmantojusi tīmekļa revolūcijas radītās iespējas: sociālos tīklus, mobilo datu pārraides tīklus un lietu internetu (<i>Internet of things</i>), kas ļauj to lietotājiem ērti dalīties ar pārtikas attēliem, receptēm, ēdiena gatavošanas video vai ierakstīt pārtikas dienasgrāmatas, veidojot plaša mēroga pārtikas datu kopas (Min et al., 2019a).
Gastrofizika	Gastrofizika – neformāls daudzdimensionāli integrētas pārtikas zinātnes termins – ir viena no jaunajām pētniecības jomām, kas pēta patērētāju pārtikas izvēli, pārtikas uztveri un daudzdimensionālu garšas pieredzi, kuras mērķis ir samazināt neveselīgu pārtikas sastāvdaļu īpatsvaru, labāk izprotot patērētāja izvēles (Spence, 2017).
Hedonistiskā pieredze	Pozitīvas afektīvas reakcijas saistībā ar pārtikas patēriņu.
Lielie dati	Lielie dati ir savākto datu kopas, kas ir tik lielas un sarežģītas, ka to apstrādei nepieciešamas jaunas tehnoloģijas, piemēram, mākslīgais intelekts. Dati var būt iegūti no daudziem un dažādiem avotiem. Bieži tie ir vienvērtīgi, piemēram, GPS sakaru dati no miljoniem mobilo tālrunu, kurus izmanto, lai mazinātu satiksmes sastrēgumu problēmu; bet var tikt izmantoti arī jaukta tipa dati, piemēram, pacientu slimības vēstures un dati par lietotņu izmantošanu. Tehnoloģijas ļauj šos datus savākt ļoti ātri, gandrīz reāllaikā, un analizēt, lai gūtu jaunas perspektīvas.
Multisensorās pieredzes	Multisensorās pieredzes ir sajūtu iespaidi, kas veido priekšstatu par apkārtējo pasauli. Koncentrējoties uz konkrētiem notikumiem, multisensorus pārdzīvojumus iegūst no tā, kā mēs redzam, dzirdam, sataustām, saģaršojam un sasmaržojam (Velasco & Obrist, 2020).
Superpatērētājs	Pārtikas izvēle ir jāskata kontekstā, un šeit autore koncentrējas uz pārtikas patērētājiem jeb superpatērētājiem (ar ko autore domā samērā turīgus un digitāli aktīvus patērētājus), kuri personificē mūsdienīgas pārmaiņas cilvēku paradumos un ikdienā. (Spence et al., 2016).
Tvīts	Mikro-emuāra ieraksts sociālajā tīklā <i>Twitter</i> , kura zīmju skaits ir 140 (līdz 2017. gadam) vai 280 (kopš 2017. gada).

Temata aktualitāte un pētījuma problēma

Pēdējos gados datorzinātņu jomā vērojama pieaugoša interese par pārtikas datu analīzi, paralēli attīstoties vairākām datorzinātņu jomām. Viena no tām ir cilvēkcentrētā datorika, kur nozīmīga ir Informācijas tehnoloģiju asociācijas (*Association for Computer Machinery, ACM*) starptautiskā cilvēka un datora saskarnes jeb mijiedarbības (HCI) konference (ACM Digital Library, n.d.), kuras ietvaros aizvien vairāk darbnīcu (Ferran et al, 2019) tiek organizētas tieši pārtikas un cilvēka mijiedarbības kontekstā, radot jaunas pieejas (UC Santa Cruz, n.d.) un skatījumu, kā datorzinātņu ietvaros analizēt pārtiku, un kāda varētu būtu pārtikas datorikas nākotne (Obrist et al., 2018). Līdzīga attīstība vērojama ACM multimediju konferencē, kur 2022. gadā notika seminārs par ēšanu, gatavošanu un ēdiena aplikācijām (CEA++, 2022). Šī datorzinātņu jomā nozīmīgā ACM konferenču attīstība pārtikas analīzes virzienā ir likumsakarīga, ņemot vērā izaicinājumus, ar kādiem sastopas gan cilvēki, gan vide, nodrošinot pārtikas sistēmu darbību visā pasaulē.

Pārtikas izvēlei un pārtikas patēriņam ir liela nozīme sabiedrības veselībā. Aptaukošanās, otrā tipa diabēts un sirds un asinsvadu slimības ir tikai dažas no veselības problēmām, kas iegūtas mūsdienu patērētāju uztura specifikas dēļ (Min et al., 2019a; Mai & Hoffmann, 2015). Katram ceturtajam pieaugušajam ir liekais svars un katrs desmitais ir aptaukojies pēc Pasaules Veselības organizācijas (PVO) datiem, kas apkopoti 2016. gadā. Tas ir saistīts ar trīskārtēju aptaukošanās izplatības pieaugumu pasaulē laikā no 1975. līdz 2016. gadam. Tādējādi aptaukošanās izplatība pasaulē uzskatāma jau par pandēmiju (Lancet Gastroenterology & Hepatology, 2021). Šo iemeslu dēļ ir svarīgi izprast pārtikas izvēles pamatā esošos faktorus un plašākus ar cilvēku pārtikas patēriņu saistītus aspektus. Viens no zināmākajiem izaicinājumiem, kas kavē pārtikas patērētājus izvēlēties veselīgākas pārtikas iespējas, ir pretrunas starp garšīgumu un veselīgumu, kam ir izšķiroša nozīme pārtikas izvēlē un kas ir svarīga aptaukošanās un citu slimību gadījumos. Šo pretrunīgumu definē kā “neveselīgs = garšīgs” (Mai & Hoffmann, 2015).

Lai gan pārtikas ietekme uz indivīda veselību ir joma, par kuru pasaules mērogā diskutē pārtikas politikas veidotāji un uztura speciālisti, ir parādījusies vēl viens jauns diskurss saistībā ar pārtikas patēriņu, proti, pārtikas ietekme uz planētas veselību, ietverot bioloģiskās daudzveidības saglabāšanu, metāna un oglekļa dioksīda rašanos ražošanas procesā, kā arī cita veida piesārņojumu rašanos, kas ietekmē un veido klimata pārmaiņas, kā arī planētas ekosistēmas kopumā (Grivins et al., 2020). Trešdaļa pasaules oglekļa dioksīda emisiju tiek radītas tieši pārtikas sistēmu ietvaros, kur lielāko pienesumu dod lauksaimniecība un zemes izmantošanas aktivitātes (aplēsts, ka tās rada ap 71% no kopējām emisijām), bet pārtikas piegādes ķēdes – transports, patēriņš, mazumtirdzniecība un citi saistītie procesi – veido attiecīgi 29% (Crippa et al., 2021). Labāk izprotot pārtikas ieradumus, var veidot tādu politiku, kas būtu vērsta gan uz cilvēku, gan planētas vajadzībām (Persson et al., 2021).

Pārtikas datorika ir starpdisciplināra joma, kas plašākā nozīmē ietver ar pārtiku saistītus pētījumus, ko veic datorzinātnēs. Izmantojot tīmekļa revolūcijas radītās iespējas – sociālos tīklus, mobilos tīklus, lietu internetu – veidojas lielapjoma datu kopas, jo jaunās tehnoloģijas ļauj to lietotājiem ērti dalīties ar pārtikas attēliem, receptēm, ēdiena gatavošanas video vai ierakstīt pārtikas dienasgrāmatas (Min et al., 2019a). Pārtikas datorikas attīstība lielā mērā veido jaunu individuālu pārtikas ieteikumu sistēmu izstrādi. Šādām pārtikas ieteikumu sistēmām, savukārt, ir jāsaprot lietotāji un ne tikai to individuālā uzvedība, bet arī sociālo grupu ietekme viņu pārtikas izvēle kontekstā. Pateicoties jauniem, uz tehnoloģijām balstītiem risinājumiem, šī starpnozaru joma ir aizvien nozīmīgāka akadēmiskajā pētniecībā, kas attiecas uz pārtikas patēriņu, sabiedrības veselības jautājumiem un arvien vairāk – arī uz planētas vides labklājību. Termins “pārtikas datorika” ietver datorzinātņu metožu – dabiskās valodas analīzes, datorredzes, mašīnmācīšanās un citu metožu – izmantošanu, analizējot dažādus ar pārtiku saistītus datus. Galvenā uzmanība tiek pievērsta liela apjoma datiem un iespējām tos apkopot un analizēt, lai izprastu pārtikas uztveri, kā arī izstrādātu ieteikumus pārtikas izvēļu prognozēšanai.

Ar liela mēroga datu izmantošanu saistās daudz cerību, lai gan joprojām trūkst regulējuma to izmantošanai. Eiropas Parlaments definē, ka liela apjoma dati jeb “lielie dati” attiecas uz tādiem liela apjoma datiem, kas nozīmīgā ātrumā tiek ģenerēti no dažādiem avotiem (Eiropas Parlaments, 2021), norādot uz vēl nebijušu mērogu (apjomu), sadrumstalotību (lielu daudzveidību) un pieaugumu reāllaikā (ātrumu). Lai gan lielo datu izmantošana aizvien vairāk tiek testēta akadēmiski, šai jomai ir plašas un vēl neizmantotas iespējas radīt jaunas pētniecības jomas ar pārtiku saistītu pētījumu nozarēs. Tā, piemēram, terminu “lielie dati” reti izmanto saistībā ar pārtikas drošību, galvenokārt tāpēc, ka pārtikas drošības dati un informācija ir izklaidēti pa pārtikas, veselības un lauksaimniecības pētniecības jomām (Marvin et al., 2017). Tātad arī šeit ir vērojami izaicinājumi, kas saistīti ar starpdisciplināritāti.

Lai gan iespējams identificēt vairākus pētījumus, kuros tiek izmantoti liela mēroga dati attiecībā uz pārtikas drošību vai citiem korektīviem aspektiem attiecībā uz pārtikas patēriņu, ir ļoti maz pētījumu, kuros galvenā uzmanība ir pievērsta pozitīviem aspektiem saistībā ar pārtikas patēriņu (Obrist et al., 2019). Tāpat Covid-19 pandēmija, kuras laikā rakstīts šis promocijas darbs, izraisījusi pētnieku papildu interesi izpētīt izmaiņas pārtikas patēriņa paradumos un praksē (Laguna et al., 2020). Mazāk pētījumu ir veikti, izmantojot plaša mēroga datus, lai noteiktu afektīvas reakcijas uz pārtiku, tostarp multisensoro pieredzi un elementus, kas ir būtiski multisensorai pārtikas uztverei (piemēram, apkārtējā vide, tekstūra, krāsa, smarža un citi aspekti). Ņemot vērā iepriekš minēto, multisensorās pieredzes definīcija ir sekojoša: “multisensorā pieredze ir sajūtu iespaidi par mūsu priekšstatu par apkārtējo pasauli. Koncentrējoties uz konkrētiem notikumiem, multisensorus pārdzīvojumus iegūst no tā, kā mēs redzam, dzirdam, sataustām, sagaršojam un sasmaržojam” (Velasco & Obrist, 2020).

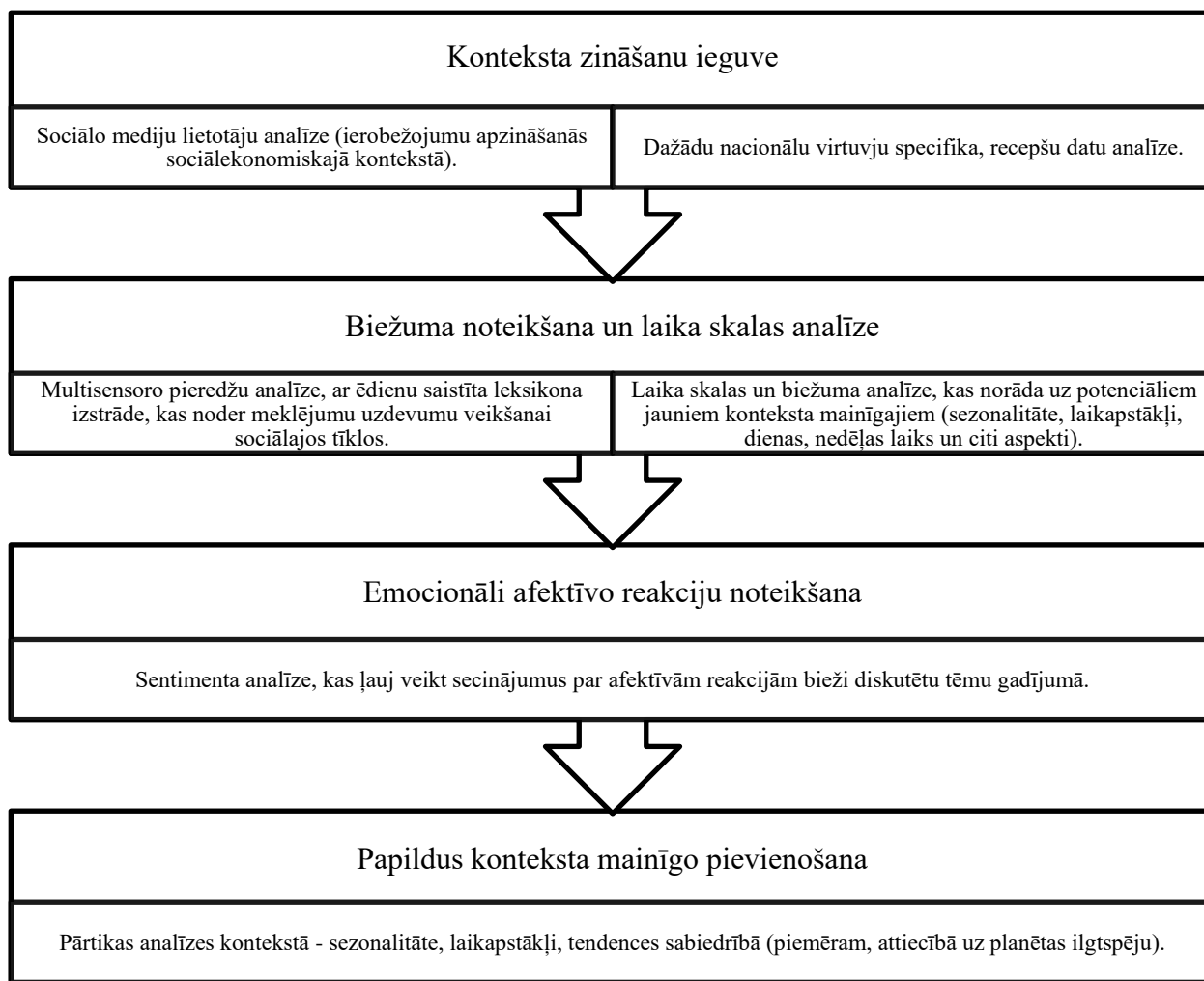
Savukārt gastrofizika, kas ir neformāls multisensoro pieredžu integrētas pārtikas zinātnes termins, ir viena no jaunajām pētniecības jomām, kas pēta patērētāju pārtikas izvēli, pārtikas uztveri un multisensoru garšas pieredzi un kuras mērķis ir samazināt neveselīgu pārtikas sastāvdaļu īpatsvaru, labāk izprotot pārtikas patērētāja uzvedību (Spence, 2017). Tā vietā, lai koncentrētos tikai uz ēdiena garšu, sastāvdaļām vai uzturvērtību, gastrofizika cenšas atklāt “visa pārējā” (*everything else*) ietekmi, kas saistīta ar ēdiena izvēli un patēriņu – vai tā būtu vide, kurā ēdiens tiek patērēts, trauki, ēdienu nosaukumi, galda piederumi utt. –, jo tas viss ietekmē degustācijas pieredzi (Spence, 2017).

Autore uzskata, ka izpratne par multisensorās pieredzes lomu un ietekmi pārtikas patēriņā ir izšķiroša sabiedrības veselības uzlabošanā. Iepriekš minētās pētniecības jomas – pārtikas datorika, gastrofizika un multisensoro pieredžu analīze kalpo kā galvenie teorētiskie pamati autores jaunajam konceptuālajam lielapjoma datu analīzes ietvaram, kura mērķis ir izprast, kā cilvēki uztver, izvēlas un patērē pārtiku.

Pētniecības joma un tās novitāte

Jauna konceptuāla ietvara pārtikas patēriņa uztveres analīzei izveide ir inovatīva un nepieciešama pieeja, kas novērš pārāk šauras pētniecības perspektīvas trūkumus un ļauj pētīt padziļināti starpdisciplināru pētniecības jomu saistībā ar pārtiku. Zemāk tabulā 1.1 īsi prezentēts konceptuālais ietvars pārtikas uztveres un patēriņa analīzei, izmantojot lielo datu pieeju.

Konceptuālais ietvars pārtikas uztveres un patēriņa analīzei, izmantojot liela mēroga datu analīzes pieeju



Turpmākajās apakšnodaļās jaunais konceptuālais ietvars tiks skaidrots un pamatots, ilustrējot katru no šī ietvara pamatelementiem ar piemēru, kas iegūti autores publikāciju kopā ar līdzautoriem, rezultātā. Tikmēr zemāk īsi prezentēta šī konceptuālā ietvara novitāte attiecībā uz datorzinātņu attīstību kopumā.

Jaunu pētniecības uzdevumu izvirzīšana: multisensoro pieredžu analīze, izmantojot lielos datus

Kamēr pārtikas datorikas pētnieki savu analītiku par pārtiku galvenokārt balsta uz pieejamajiem lielajiem pārtikas datiem, piemēram, receptēm, ēdienu attēliem, kognitīvo zinātņu pētnieki mēdz aplūkot plašākus ar pārtikas patēriņu saistītus jautājumus (Kāle & Agbozo, 2021). Līdz šim pētījumi par “visa cita” nozīmi liela mēroga datos, analizējot pārtikas izpratni un patēriņa jautājumus, ir bijuši reti, tāpēc multisensoru pieredžu un sentimenta analīze no *Twitter* lietotāju datiem, ir nozīmīgs pirmais solis, lai ilustrētu iespējas, kādas paveras, analizējot lielos datus pārtikas izpratnes un patēriņa nozīmes kontekstos.

Sabiedrības tendenču un ietekmes uz indivīdu analīze, nodrošinot jaunas zināšanas personalizēto pārtikas (un veselības) lietotņu izstrādātājiem

Ņemot vērā, ka viens no nozīmīgākajiem attīstības virzieniem pārtikas datorikā ir personalizētu

lietotņu attīstība, kas apkopo lielu apjomu par pārtiku pieejamos datus un pielāgo to indivīda vajadzībām, autores darbs dod ieguldījumu arī šīs jomas attīstībā, jo ieteikumu uzdevuma izpildei var noderēt prognoze no sociālajiem medijiem. Piemēram, “no sociālajiem medijiem prognozētās lietotāju pārtikas preferences ir nozīmīgs solis, lai attīstītu personalizētus ieteikumus par pārtiku” (Min et al., 2019a). Jaunais konceptuālais ietvars palīdz analizēt ar pārtikas patēriņu saistītas tendences sabiedrībā un var palīdzēt izprast indivīda un grupu uzvedības sasaisti attiecībā uz pārtikas patēriņu, analizējot lielos datus.

Inovatīvu metodoloģisku formātu veidošana, tādējādi veicinot pārtikas datorikas izaugsmi un starpdisciplināritāti pētniecībā kā tādu

Tā kā pārtikas datorika ir jauna pētniecības joma, kas strauji attīstās, šī promocijas darba izstrāde atbilst vajadzībām, ko jomas zinātnieki ir pauduši (Nag et al., 2017), un tās ir jaunu pieeju un jaunu konceptuālu ietvaru radīšana, lai pēc iespējas labāk izmantotu tehnoloģiju un lielo datu sniegtās iespējas cilvēka un pārtikas analizē, metodoloģiju attīstībā un jaunu analītikas pieeju veidošanā. Kopumā iztrūkst metodoloģiju par to, kā ietvert lielos datus par pārtiku analizē, un jaunu pieeju veidošanās ir pašā sākumā (Jin et al., 2020), tāpēc autores darbs ir nozīmīgs darbs pārtikas datorikas jomas attīstībā kopumā.

Pētījuma motivācija, uzdevumi un mērķi

Šī promocijas darba kopsavilkuma uzdevums ir izveidot jaunu konceptuālu ietvaru, kā iekļaut liela mēroga datus – šajā gadījumā pārtikas uztveres un pārtikas patēriņa datus – analizē, kā arī ilustrēt šī konceptuālā ietvara darbību, veidojot uz jaunām metodēm bāzētu rakstu krājumu pārtikas datorikas jomā. Līdz ar to tiek piedāvāts ne tikai teorētisks ietvars, kas kalpo par pamatu pētniecības jautājumu izstrādei, bet autore ar savām un līdzautoru publikācijām ilustrē šī konceptuālā ietvara ieviešanu arī praksē, kas attiecas uz jaunu, inovatīvu zinātnisko rakstu izstrādi.

Pamatojoties uz novērojumiem, ka pasaulē, kurā datu kļūst arvien vairāk, joprojām trūkst metodiskas pieejas šo datu analīzei, autores motivācija ir izveidot jaunu konceptuālu pētniecības ietvaru, kas iekļauj atšķirīgās pieejas no pārtikas datorikas, kognitīvās un sociālās zinātnes par pārtikas uztveri un pārtikas patēriņu. Izpratne par starpdisciplināritātes nepieciešamību pastāv dažādos pārvaldības līmeņos (Allmendinger, 2015) taču vienlaikus tiek atzīts, ka starpdisciplināri pētījumi joprojām ir problemātiski. Tomēr tieši datorzinātņu joma ir tā, kur starpdisciplināritāti ir iespējams veicināt, un tādas jomas kā pārtikas datorikas attīstība, tam ir piemērs.

Visbeidzot, šīs promocijas darba kopsavilkuma mērķis ir ilustrēt mazo valodu – tādu, kā latviešu valodas – datu lietderību, izmantojot Latvijas *Twitter* Ēdāju Korpusu (*Latvian Twitter Eater Corpus* (LTEC)) – unikālu resursu, kas paredzēts ar pārtiku saistītu latviešu valodas tvītu analīzei. Tas varētu kalpot par piemēru citām mazajām valodām un veicināt labāku izpratni par atšķirībām pārtikas aprakstos atkarībā no lietotās valodas (Fenko et al., 2010). Kā liecina pētījumi, pārtika lielā mērā ir saistīta ar kultūru un klimatu, kā arī valodu, kas atklāj konkrētas īpatnības, kas jāņem vērā, veidojot pārtikas sistēmu politiku un ieteikumus labākiem sabiedrības pārtikas patēriņa paradumiem (Jurafsky, 2015; Metcalfe, 2019).

Pētniecības jautājums

Pamatojoties uz pētniecības motivāciju, mērķiem un uzdevumiem, šajā promocijas darbā tiek atbildēts uz sekojošo pētniecības jautājumu: **kādas metodes, izmantojot liela mēroga datus, pielietot, lai labāk izprastu pārtikas patērētājus?**

Atbildot uz šo pētniecības jautājumu, autores mērķi ir:

- 1) izstrādāt jaunu starpdisciplināru pētniecības ietvaru, kas veicina starpdisciplināritātes izpratni un metodoloģijas attīstību, apvienojot pārtikas datorikas un kognitīvo zinātņu jomas;
- 2) sniegt piemēru par jaunā pētniecības ietvara darbību, izstrādājot metodoloģijas un publicējot zinātniskos rakstus, kas ietver multisensoro pieredzi un pārtikas uztveres analīzi, kas balstīta uz lielajiem datiem;

- 3) dot būtisku ieguldījumu pārtikas datorikas kā jomas attīstībā, ilustrējot pētniecības iespējas, analizējot sociālo tīklu (konkrēti *Twitter*) ietvaros radītos lielos datus;
- 4) ieteikt turpmākos pētniecības virzienus, kas varētu būt noderīgi akadēmiskajai sabiedrībai, kura strādā pārtikas patērētāju pētniecības jomā un uztveres pētījumos;
- 5) sniegt ieguldījumu sabiedrības veselības politikas veidošanā, veicinot uz datiem balstītu pieeju, lai panāktu veselīgāku pārtikas patēriņu sabiedrībā un tādējādi augstāku dzīves kvalitāti pārtikas patērētājiem.

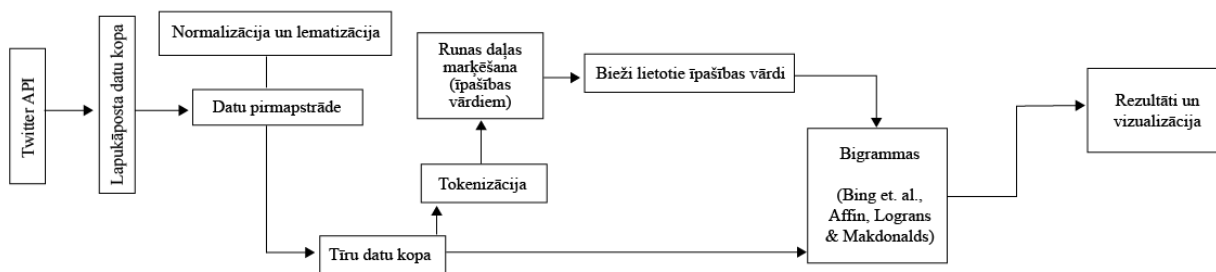
Pētniecības metodes

Autore savā darbā izmantojusi vairākas pētniecības metodes, kas īsi aprakstītas šajā nodaļā. Tomēr, ņemot vērā, ka katrā no autores publikācijām tika izmantotas dažādas pētniecības metodes, tad ar tām pilnībā iespējams iepazīties, izlasot rakstu krājumu.

Lai noteiktu būtiskākos ar sabiedrības veselības un pārtikas patēriņu saistītus jautājumus, tika analizētas datorzinātņu, sociālo un kognitīvo zinātņu publikācijas. Dziļākam ieskatam, kas saistīti ar ēdiena izvēli, uztveri un patēriņu, autore analizēja literatūru, kas saistīta ar pārtikas datoriku, gastrofiziku, multisensoro pieredzi, kā arī citus ar veselību un ēdiena uztveri saistītus pētījumus. Padziļināti tika analizēta ar pārtikas datoriku saistītā literatūra – analītiskās metodes, datu kopu daudzveidība un pieejamība.

Tālāk autores darbā izmantotās metodes var sadalīt atbilstoši datiem, kas tika analizēti. Tos var sadalīt divās daļās – recepšu dati un sociālo mediju dati. Recepšu dati tika analizēti ar tēmu modelēšanas (*topic modeling*) metodoloģiju. Tēmu modelēšana ir neuzraudzīts modelis (*unsupervised model*), kas apgūst pamattēmu kopumu. **Tēmu modelēšana** tika izmantota, lai iegūtu informāciju par noteicošajām tēmām konkrētu zemju recepšu kopumā, kam sekoja padziļinātāka analīze (Nikolenko et al., 2017). Tika izmantots latentais Dirihlē modelis (latent Dirichlet allocation (LDA)). Sākotnējā datu kopa, kas atvasināta no Yummly-66 k datu kopas (sastāv no 66 615 receptēm no desmit virtuvēm vietnē *Yummly* – mājaslapas, kas ģenerē personalizētas receptes un kalpo kā recepšu meklētājs) sagatavo ēdienu recepšu krātuvi. Pēc tam tēmas modeļa funkcija tika veidota, pamatojoties uz LDA (latent Dirichlet sadalījuma) modeli, lai iegūtu desmit labākos (10) tēmu modeļus no korpusa. LDA ir varbūtības modelis, kas izskaidro vārdu līdzaspastāvēšanu tekstā, kas izriet no relatīvi neliela iespējamo semantisko grupu vai tēmu skaita (Ihler & Newman, 2011). LDA izmanto pētnieki dokumentu klasificēšanai, rakstu atklāšanai un tendencēm – piemēram, medicīnas jomā (Tran et al., 2019), un autores un līdzautoru darbā tiek pierādīts, ka tēmu modelēšana ir noderīga arī pārtikas datorikas jautājumu analīzē (Kāle un Agbozo, 2021).

Pielietotās metodes, izmantojot sociālo tīklu lielos datus varam apskatīt attiecībā uz to, kādas valodas lielie dati ir izmantoti. Angļu valodas *Twitter* datiem vispirms bija svarīgi pielietot **teksta ieguves jeb tekstizraces** (*text mining*) metodoloģiju. Teksizrace tika izmantota zināšanu ieguvei kā “datu analīzes process, lai iegūtu galvenos jēdzienus un tēmas un atklātu slēptās attiecības un tendences” (Khatai et al., 2019; Lucini et al., 2017). **Bigrammu ekstrakcija** tika izmantota, lai no datu kopas iegūtu saturīgu kontekstuālu informāciju. N-gramas (vai bigrammas) ir N rakstzīmju secība (Gargouri et al., 2003). Tika veikti izvēlētu atslēgas vārdu meklējumi un n-gramas (bigrammas) noteiktas, analizējot lietvārdus un īpašības vārdus. Lograna-Makdonalda (Loughran-McDonald) leksikons un AFFIN semantiskais leksikons izmantots, lai klasificētu pozitīvus un negatīvus tvīta sentimentus, un noteiktu visbiežākās bigrammas (īpašības vārdiem un lietvārdiem) (Kāle & Agbozo, 2020b). Grafiski šis process vizualizēts attēlā 1.1. zemāk.



Att. 1.1: Pētniecības metode (Kāle un Agbozo, 2020b)

Savukārt, latviešu valodas *Twitter* teksta datu analīzei tika pielietotas citas metodes. Analīzei izmantojām nesen publicētu tvītu datu kopu – latviešu *Twitter* Eater Corpus (LTEC – Sproģis & Rikters, 2020). Datu kopa veidota deviņu gadu laikā ar *Twitter* straumēšanas palīdzību - API (*Twitter* izstrādātāju platforma n.d.a: n.pag.), vadoties pēc 363 ar ēšanu saistītiem atslēgvārdiem. Datu kopas kopējais lielums pārsniedz 2,3 milj. tvītu, ko ģenerējuši kopumā 168 703 lietotāji. Katrā tvītā LTEC ir primārie lauki, piemēram, *Twitter* ID, pats teksts, autors un tā tapšanas laiks. Atkarībā no teksta, metadatiem var būt arī citas neobligātas vērtības, aprakstot pievienotās multivides datnes, atrašanās vietu un jebkuru pārtikas produktu vai dzērienu, kas minēts tvītu tekstā. **Sentimenta analīzes pieeja** tika pielāgota un izmantota latviešu valodas tvītiem (Thakar & Pinnis, 2020) kopā ar latviešu valodai raksturīgu modeli (Vīksna un Skadiņa, 2020), kas tika pielietots visā LTEC. Šāda pieeja ļāva sentimenta klasifikatoram atšķirt negatīvus, pozitīvus un neitrālus tvītus LTEC sentimenta apmācības datu apakškopā.

Rezultātu aprobēšana

Tālāk seko autores un līdzautoru publikāciju un konferenču prezentāciju saraksts. Autores ieguldījuma līmenis katrā publikācijā ir 60% un tas ir noteikts, vienojoties ar līdzautoriem.

Publikāciju saraksts

1. Kāle, M., Šķilters, J. & Rikters, M. (2021). ‘Tracing multisensory food experiences on *Twitter*’, *International Journal of Food Design*, 6:2, pp. 181–212.
2. Kāle, M. & Agbozo, E. (2021) ‘Utility of large-scale recipe data in food computing.’ *Baltic Journal of Modern Computing*, 9(2), 155-165.
3. Kāle, M. & Rikters, M. (2021). “Fragmented and Valuable: Following Sentiment Changes in Food Tweets”, in workshop proceedings “*Smell, Taste, and Temperature Interfaces (STT21)*”; *CHI 2021 Workshop*.
4. Kāle, M. & Agbozo, E. (2020). “Healthy Food Depiction on Social Media: The Case of Kale on *Twitter*”. *DHN Post-Proceedings*.
5. Kāle, M. & Agbozo, E. (2020). “Tracing complexity in food blogging entries” *CEUR Workshop Proceedings*.

Referāti konferencēs

1. Dalība panelīdiskusijā “Pārtikas datorikas nākotne”. *Informācijas tehnoloģiju asociācijas darbnīca par veselības un pārtikas jautājumiem (MaDiMa) un starptautisks seminārs par multivīdi ēdiena gatavošanai, ēšanai un saistītajām lietotnēm (CEA++) (30th ACM International Conference on Multimedia)*. 10.10.2022., Lisabona, Portugāle.
2. Prezentācija “Multisensorā pārtika un sociālie mediji”. *Pārtikas datorikas darbnīca, (Simula Research Laboratory)*, 21.06.2022. Oslo, Norvēģija.

3. Informācijas tehnoloģiju asociācijas darbnīca datoru un cilvēku mijiedarbības jautājumos (ACM SIGCHI) ikgadējā konference – ACM konference par cilvēkfaktoriem datorsistēmās (CHI 2021); tiešsaistē, 2021. gada 8.–13. maijā.
 - b) Autore pētījumu prezentēja divās darbnīcās:
 - i) STT21 – smaržu, garšu un temperatūras saskarņu jeb mijiedarbības darbnīca. Prezentācija: “Fragmentāri un vērtīgi: sentimenta izmaiņu noteikšana pārtikas tvītos”.
 - ii) Cilvēka un pārtikas mijiedarbības nākotne. Prezentācija: “Gaļas nākotne – tvītu sentimentu analīze”.
4. Prezentācija “Sentimenta izmaiņu izsekošana pārtikas tvītos” *Latvijas Universitātes 79. starptautiskā zinātniskā konference; sesija “Kognitīvā zinātne”*, 12.02.2021. Rīga, Latvija.
5. Prezentācija “Sarežģītības izsekošana pārtikas emuāru ierakstos” *DHN2020 starptautiskā konference par digitālajām humanitārajām tēmām*. 22.10.2020., tiešsaistē.
6. Prezentācija “Kognitīvās zinātnes jēdzienu derīgums, pielietojot kvantitatīvas datu analīzes metodes: pārtikas datorikas gadījuma analīze” *Latvijas Universitātes 78. starptautiskā zinātniskā konference; sesija “Kognitīvā zinātne”* 14.02.2020. Rīga, Latvija.
7. Prezentācija “Recepšu datu analīze: pārtikas datorikas un kognitīvo zinātņu sasaiste” *Latvijas Universitātes 14. starptautiskais izziņas, loģikas un komunikācijas simpozijs. Maņu sasaiste: modalitātes uztvere* 7.12.2019. Rīga, Latvija.

Galvenie rezultāti

Autores darba rezultātā tika izveidots jauns konceptuāls ietvars pārtikas uztveres un patēriņa analīzei, izmantojot lielo datu sniegtās iespējas un priekšrocības, tādējādi attīstot inovatīvu pieeju, kas iespējama, tikai veicot patiesi starpdisciplināru pētījumu. Jaunais konceptuālais ietvars sastāv no četriem galvenajiem blokiem:

1. konteksta zināšanu iegūšana par tēmu;
2. tēmas biežuma konstatēšana (*detecting frequency rates*);
3. afektīvu reakciju noteikšana un laika grafika analīze;
4. papildu konteksta mainīgo pievienošana.

Lai gan rezultāti ir detalizēti izklāstīti vēlāk, zemāk ir kopsavilkums ar galvenajiem promocijas darba rezultātiem, lai gūtu pārskatu par secinājumiem, kā arī metodēm, kas tika izmantotas rezultātu sasniegšanai.

Pielietotās metodes konteksta zināšanu iegūšanai par tēmu: tēmu modelēšana, datorzinātņu, sociālo un kognitīvo zinātņu literatūras avotu izpēte. Rezultāti:

1. Izpratne par veselīgu pārtiku ir atkarīga no konteksta, kādā tiek patērēts ēdiens, konkrētā patērētāja veselības stāvokļa, kā arī citiem sociālajiem, ekonomiskajiem un vides faktoriem, kas apskatīti šajā kopsavilkumā.
2. Viena no šī darba centrālajām tēmām ir “neveselīgs = garšīgs” intuīcija (Mai et al., 2011), kas rada kognitīvu disonansi pārtikas patērētājiem, kam veselīga pārtika uzreiz asociējas ar negaršīgāku pārtiku, kamēr neveselīgi ēdieni un to garšas īpašības tiek aprakstītas daiļrunīgi un sensuāli (Jurafsky, 2015). Tieši šādus rezultātus, kur veselīgai pārtikai piedēvētas drīzāk pragmatiski un racionāli labumi, nevis garšīgums un patika to patērējot, varam vērot arī sociālo tīklu, konkrēti *Twitter*, analizē (Kāle & Agbozo, 2020a; Kāle & Agbozo, 2020b).
3. Patērētāji, kas aktīvi darbojas sociālajos medijos, galvenokārt ir turīga, digitāli saistīta sabiedrības daļa, līdz ar to no sociālo mediju datiem nav iespējams izdarīt secinājumus par tiem, kas dzīvo nabadzībā un kuriem ir zema digitālo mediju pratība.

Pielietotās metodes tēmas biežuma konstatēšanai: tekstizrace un bigrammu analīze. Rezultāti:

1. Analizējot dažādu tēmu biežumu, kas parādās sociālajos tīklos, iespējams iegūt zināšanas par diskursiem, kas dominē sabiedrībā saistībā ar, piemēram, veselīgu pārtiku vai garšīgu pārtiku.
2. Biežuma noteikšanu sociālo tīklu platformās iespējams piemērot konkrētiem pārtikas produktiem vai pārtikas produktu grupām (piemēram, vegānu pārtika, noteikti populāri pārtikas produkti konkrētajās sabiedrībās un citas ar veselību un pārtiku saistītas tēmas – glutēna panesamība, piemēram).
3. Tāpat šādā veidā var iegūt vārdu krājumu, ko cilvēki izmanto noteiktu pārtikas produktu vai pārtikas patērēšanas pieredzes aspektu aprakstīšanai, kas ir specifiski katrai valodai, un var noderēt dažādu citu pētniecības jautājumu apskatīšanai par to, kā cilvēki konkrētās sabiedrībās runā par ēdienu un pārtiku kopumā.

Pielietotās metodes afektīvu reakciju un laika grafika noteikšanai: sentimenta analīze. Rezultāti:

1. Spēja izsekot afektīvajām reakcijām attiecībā uz pārtiku, tāpat arī laika grafikam (kurā dienas daļā cilvēki sociālos medijos dalās ar konkrēto informāciju), ļauj padziļināti izprast sociālo tīklu ietekmi uz pārtikas patērētājiem un ar to saistītās emocijas.
2. Tāpat šī pieeja dod ieguldījumu dažādu datorzinātņu jomu lietotņu izstrādē – gan personalizētu veselības uzraudzībā, gan pārtikas patēriņa izpratnes veidošanai par afektīvo reakciju dinamiku, un ietekmi uz pārtikas izvēli un patēriņu.

Pielietotās metodes papildu konteksta mainīgo pievienošanai: sentimenta analīze par jauniem konteksta mainīgajiem. Rezultāti:

1. Zināšanas par sezonālajām pārtikas patēriņa tendencēm, atbilstoši nacionālajai kultūrai un klimatam, ļauj precīzāk prognozēt pārtikas patēriņu, un līdz ar to labāk arī pārvaldīt resursus, piemēram, savlaicīgi sasaistot laikapstākļu datus ar potenciālajiem patēriņa datiem.
2. Kopumā pētījumi par laikapstākļu ietekmi uz sabiedrības noskaņojumu un patēriņa tendencēm ir visai reti, taču atklāj nozīmīgu faktoru, kas būtu jāņem vērā. Šāds pētniecības jautājums par laikapstākļu ietekmi uz pārtikas uztveri un patēriņu ir bijis iespējams tikai pateicoties iepriekšējiem pētījumiem kognitīvās zinātnes jomā, kur novitāte ir bijusi datorzinātņu metodoloģijas piemērošana, lai atbildētu uz šādu pētniecības jautājumu.

Ar šo jauno konceptuālo ietvaru autores mērķis ir novērst nepilnības mono-disciplinārās pētniecības paradīgumā, piedāvājot formātus, kurus var izmantot starpdisciplinārai pārtikas patēriņa analīzei, izmantojot aizvien vairāk un plašāk pieejamos lielos datus (Eiropas Parlaments, 2021).

Praktiskais pielietojums

Tā kā pārtikas pētniecība ir visai starpdisciplināra, tā var kalpot par piemēru tādu starpnozaru jautājumu risināšanai, kur vienas nozares pētniecība nevar sniegt pietiekami precīzas un izsmeļošas atbildes. No praktiskās pielietojamības aspekta, šie rezultāti varētu būt noderīgi politikas veidotāju darbā, kur patērētāju uzmanība tiek vērtā gan personīgās veselības, gan planētas labklājības un ilgtspējas virzienā, lai risinātu iedzīvotāju skaita pieaugumu un aizvien pieaugošās vajadzības pēc pārtikas (World Health Organization, 2021). Pārtikas datorikas kā datorzinātņu jomas attīstība ir tieši saistīta ar ilgtspējas izaicinājumiem un iespējam, ko digitalizācija, tehnoloģijas un lieli dati var sniegt, lai šos izaicinājumus risinātu, un to uzsver visos nozīmīgākajos pasaules ekonomikas forumos (World Economic Forum, 2020) un diskutē A līmeņa konferencēs visā pasaulē (Yamakata et al., 2022).

Autore un publikāciju līdzautori jau saņēmuši atsauksmes no citiem pētniekiem par šādu pieeju: “šo novatorisko rīku, kas analizē iedzīvotāju noskaņojumu attiecībā uz populāriem pārtikas produktiem, nedēļas dienu ietvarā, varētu izmantot, piemēram, lai korelētu ēšanas paradumus, kolektīvos uzskatus un tirgus ietekmes. Tas palīdzētu izstrādāt efektīvākas stratēģijas, vai nu izmantojot reklāmu, vai politikas plānošanu

un izglītošanas pieeju, lai atbalstītu pārtikas patērētāju klimatom draudzīgas izvēles (piemēram, pārtikas patēriņu ar mazākām oglekļa emisijām). Tāpat šī analīze varētu būt noderīga pārtikas nozares speciālistiem, kā arī pārtikas tirgotājiem (piemēram, zināšanām par to, kā sabiedrībā atbalsojas ar pārtiku saistīti vēstījumi)” (Velasco et al., 2021).

Izstrādātais un prezentētais jaunais konceptuālais ietvars pārtikas patēriņa analīzei, izmantojot lielos datus, var kalpot par piemēru citu ar uztveri saistītu pētījumu izstrādei, piemēram, dažādu ar veselību saistītu faktoru uztveres pētījumiem. Pēc šī jaunā konceptuālā ietvara parauga var veikt analīzi par veselīguma uztveri sabiedrībā. Veselība ir joma, kur plaši sastopami viltus fakti, dažādi mīti un stāsti par alternatīvās medicīnas sistēmām. Analizējot sabiedrības izpratni par dažādiem ar veselību saistītiem jautājumiem, iespējams izstrādāt labākas politikas atbildes, kas īpaši noderētu dažādu pandēmiju laikā. Visas šīs zināšanas ir būtiskas, izstrādājot daļēji automatizētus vai automatizētus risinājumus ar indivīda veselību un ilgtspēju kopumā saistītus jautājumus.

Tāpat arī rezultāti var būt noderīgi tiem pārtikas datorikas pētniekiem, kas pievērsušies personalizētu pārtikas un veselības lietotņu izstrādei (Rostami et al., 2020). Konteksta zināšanas par sabiedrību, kur šīs aplikācijas ieviest un metodoloģija lielo datu izmantošanai analītikā, ir nozīmīgs pienesums. Arvien vairāk tiek pētīts arī mūsu analogo un digitālo dzīvju saistība un to, kā pieredze digitālajā dzīvē (tai skaitā sociālajos tīklos) ietekmē dažādas mūsu izvēles analogajā pasaulē. Autores un līdzautoru publikācijas var kalpot kā nozīmīgs datu avots tiem, kas vēlas pētīt digitālās un analogās dzīves sasaisti, līdzīgi kā publikācijā “Kā digitālā pārtika ietekmē mūsu analogo dzīvi: pārtikas fotografēšanas ietekme uz veselīgas ēšanas paradumiem” (Andersen et al., 2021), kā arī citus ar multimediju jomu datorzinātnēs saistītus jautājumus (Yamakata et al., 2022).

1. TEORĒTISKAIS PAMATOJUMS

Savās publikācijās autore un līdzautori uzdod jaunus pētniecības jautājumus, kas līdz šim nav analizēti ar datorikas metodēm, izveidojot jaunu konceptuālu ietvaru analīzei, kas ietver:

1. konteksta zināšanu iegūšanu par tēmu;
2. tēmas biežuma konstatēšanu;
3. afektīvu reakciju noteikšanu un laika grafika analīzi;
4. papildu konteksta mainīgo pievienošanu.

Lai iepazīstinātu ar pamatojumu šo pētniecības jautājumu un kontekstuālā ietvara bloku izstrādē, šajā nodaļā tiks izvērsti teorētiskais pamatojums, kam sekos pētniecības metožu apraksts, galvenie rezultāti, kā arī ierobežojumi, ar kuriem autore sastapsies savā darbā.

1.1 Pārtikas uztveres un multisensoro pieredžu pētījumi

Kopumā pārtikas jomu kā tādu raksturo daudz neefektivitāšu – vai tā būtu veselīgas pārtikas izvēle vai pārtikas atkritumu pārvaldība – attīstītajās pasaules valstīs pārtikas atkritumi ir līdz pat 40%, kur lielākā daļa rodas tieši mājsaimniecību pārtikas iegādes paradumu dēļ. Pārtikas izvēle bieži ir impulsīva, to ietekmē apkārtējā vide un sabiedrības normas, līdz ar to varam secināt, ka pārtikas izvēle ir kontekstuāla: tās pieejamība, kultūras un sabiedrības, kā arī personīgie ieradumi – tas viss ietekmē ar pārtiku saistītu lēmumu pieņemšanu. Piedevām šie ietekmējošie faktori ir neprognozējami un ir mainīgi. Patērētāju izvēle attiecībā uz pārtiku ne vienmēr pārtikas izvēle ir racionāli pamatojama – bieži pārsvaru pār izvēles procesu ņem emocijas un dažādi nejauši konteksta faktori (Kāle & Agbozo, 2020).

Daudz cerību attiecībā uz pārtikas sistēmu neefektivitāšu risinājumiem liktas uz datorzinātņu attīstību, kad ar jauniem tehnoloģiskiem risinājumiem būtu iespējams apskatīt gan dažādus riskus lauksaimniecībā, gan pārtikas piegādes ķēdēs un visbeidzot – paša patērētāja pārtikas izvēles. Vienlaikus strauji attīstās gan precīzā lauksaimniecība (Carolan, 2015), gan dažādi personalizētu pārtikas lietotņu risinājumi (Nag et al., 2021). Un tomēr, lai arī gaidas ir augstas, informāciju un komunikāciju tehnoloģiju (IKT) sektors joprojām saskaras ar nozīmīgiem izaicinājumiem, tieši attiecībā uz ilgtspējas risinājumu izstrādi pārtikas sistēmām (Svenfelt & Zapico, 2016). Autore uzskata, ka izaicinājumus rada tieši starpdisciplināritātes jautājumi, un līdz ar to šis darbs uzskatāms par paraugu, kā iespējams strādāt pie ilgtspējas un veselības jautājumiem un ar datorzinātņu jomas metodoloģiju palīdzību iegūt jaunas, vērtīgas zināšanas.

Jāņem vērā, ka autore analizē tādu pārtikas patērētāju, ko var dēvēt par superpatērētāju (*super-consumer*), ar ko autore domā samērā turīgu un digitāli aktīvu patērētāju, kurš personificē pārmaiņas cilvēka ikdienā un paradumos (Spence et al., 2016). Autore apzinās, ka lielu daļu pasaules sabiedrības nevar uzskatīt par superpatērētājiem ierobežotās pārtikas pieejamības dēļ. Sabiedrība kļūst aizvien polarizētāka, kurā līdzās superpatērētājiem pastāv aizvien pieaugoša nabadzība, un paradoksāli tieši pārtikas jomā strādājošie nereti saņem vismazāko atalgojumu un ir finansiāli un sociāli visapdraudētākie (Grivins et al., 2020). Dziļāka izpratne par tiem, kas aktīvi piedalās sociālo tīklu diskusijās, ir nozīmīga plašākam datorzinātņu pētnieku lokam, kas strādā ar sociālo tīklu ierakstu analīzi – svarīgi apzināties, ka dati, kas rodas ierakstu un fotogrāfiju ievietošanu rezultātā sociālajos tīklos, primāri ir radušies salīdzinoši labklājīgās sabiedrības grupās.

Ņemot vērā grūtības noteikt faktorus, kas ietekmē superpatērētāja pārtikas preferences, autore secina, ka lielākā daļa pētījumu līdz šim ir balstījušies uz koriģējošiem pārtikas aspektiem – tādiem kā svara samazināšana, veselīgākas ēšanas paradumu izpēte u.c. (Obrist et al., 2019). Tomēr līdztekus jāņem vērā, ka ik dienu pieaug tekstuāla un vizuāla informācija par pārtiku un aizvien lielāka nozīme digitālajos medijos ir tādiem datiem par pārtiku, kas ietver slavenību blogus, dažādus TV šovus ar gatavošanas elementiem un citiem izklaidējošiem ar pārtiku saistītiem pasākumiem (Spence et al., 2016). Liela nozīme ir hedonistiskās pieredzes pārnesei uz jaunajiem medijiem, kas kļūst aizvien nozīmīgāka arī pārtikas piegādātāju un patērētāju komunikācijā, un kas aizvien intensīvāk norit gan digitālajā, gan analogajā vidē (Andersen et al.,

2021).

Viens no jēdzieniem, ko izmanto, lai noteiktu pozitīvo ar pārtiku saistīto pieredzi, ir sarežģītības (*complexity*) jēdziens, kas tiek saprasts kā ar pārtikas patēriņu saistītās pozitīvās pieredzes daudzveidība un bagātība, un, neraugoties uz šī jēdziena neskaidro definīciju, tas ļauj diferencēt pārtikas patēriņa pieredzi no vienkāršotas un racionālas pārtikas izvēles perspektīvas (Kāle un Agbozo, 2020a). Šī teorētiskā pieeja ir novatoriska attiecībā uz plaša mēroga sociālo tīklu datu analīzi, kas saistīta ar pārtiku un ļauj izsekot pārtikas uztverei, kas atspoguļo hedonistisku pieredzi. Šeit sarežģītība tiek analizēta ar bigrammu palīdzību, apskatot lietvārdus un īpašības vārdus, kas raksturo konkrētu pārtikas produktu. Sīkāk šī pieeja tiks paskaidrota metodoloģijas un rezultātu analīzes sadaļās. Tomēr viens no galvenajiem secinājumiem ir tas, ka veselīgas pārtikas aprakstiem pietrūkst hedonistisku noskaņu salīdzinājumā ar komforta ēdieniem, piemēram, šokolādi un vīnu (Kāle & Agbozo; 2020a; Kāle & Agbozo 2020b). Kamēr vīnu dēvē par “plecīgu”, “muskuļotu” un “grēcīgu” un desertu par “seksiģu” un “pavedinošu” (Jurafsky, 2015), veselīgu pārtiku raksturo tās pragmatiskās īpašības, piemēram, “viegli pagatavojama” un “veselīga”. Veselīgai pārtikai trūkst daiļrunīgo komforta ēdienu aprakstu, līdz ar to – neveidojas asociācijas, ka veselīga pārtika var būt arī garšīga. Tas, savukārt, nozīmē to, ka vēloties sevi iepriecināt, cilvēki izvēlēsies pārtiku, kas jau sākotnēji nāk komplektā ar daiļrunīgu, iztēli rosinošu garšas aprakstu, nevis pragmatisku, racionālu pieeju savas veselības labā.

Multisensorās pieredzes pētniecība noder, lai izpētītu dažādus aspektus, kas raksturo patērētāja afektīvi emocionālo attieksmi pret pārtiku. Kopumā multisensorā pieredze veidojas no tā, kā mēs uztveram apkārtējo pasauli, kuras centrā ir konkrēti notikumi, kā rezultātā rodas iespaidi ar sensoriem elementiem tajos (Velasco & Obrist, 2020). Pārtikas pētniecības joma ir īpaši labvēlīga multisensoras pieredzes izpētei, jo tā nodrošina tādus ar patēriņu saistītus elementus (piemēram, restorānu apmeklējums, ēdiena vizuālais noformējums u.c.), kas ir būtiski multisensoras pieredzes izsekošanai un pārtikas datorikas attīstībai. Jēdziens “multisensorās pieredzes” ietver gan ēdienam raksturīgos sensoriskos faktoros (piemēram, krāsu, smaržu, tekstūru, viskozitāti), gan ārējos faktoros, kas saistīti ar produktu iepakojumu, galda piederumiem, kontekstuāliem faktoriem, piemēram, fona mūziku, atmosfēras sajūtu u.c., kas nosaka to, kā mēs uztveram ēdienu (Wang et al., 2019). Autore kopā ar līdzautoriem izveidojusi pieeju daudznozaru pārtikas pieredzes analīzei, tādai kā multisensoro pieredžu analīzei, izmantojot pārtikas datorikas metodes – lielo datu analīzi ar esošajām datorzinātņu metodēm dabiskās valodas analīzei.

1.2. Pārtikas datorika un dabiskās valodas dati

Pārtikas datorika ir pētniecības virziens, kas saistīts ar pārtiku, izmantojot datorzinātņu metodoloģiju. Tā ir salīdzinoši jauna datorzinātņu apakšnozare, kas strauji attīstījusies kopš 2017. gada, un sastāv no pieciem pamatuzdevumiem: pārtikas datu uztveres, atpazīšanas, atgūšanas un ieteikumu veidošanas, kā arī prognozēšanas un uzraudzības (Min et al., 2019a). Šie pētniecības uzdevumi izriet no lielu ar pārtiku saistītu datu kopu pieejamības: sākot ar attēliem, attēlu atpazīšanu, datorredzes risinājumiem, dabiskās valodas datu analīzi un beidzot ar personalizētu lietotņu izstrādi. Sagaidāms, ka pārtikas datorika strauji attīstīsies personalizētu veselības lietotņu virzienā, uz ko norāda iestrādnes un attīstība industrijā (skat. – MediPiatto (Vasiloglou et al., 2020) un *JoinZoe*, kas, izmantojot cilvēka mikrobioma datus no PREDICT pētījumiem (The PREDICT program, 2020), izstrādāja algoritmus personalizētai pārtikas ieteikumu sistēmai, veiksmīgi piesaistot vairākus desmitus miljonu ASV dolāru (Personalized nutrition startup Zoe closes out Series B at \$53M total raise, 2021)).

No iepriekš uzskaitītajām dažādu veidu datu kopām šajā pētījumā autore izmantoja dabiskās valodas datus – no receptēm un tvītiem sociālajos tīklos, konkrēti *Twitter*. Autore pētījusi dabiskās valodas datus un veidu, kā mēs varam analizēt vārdu nozīmi, to kontekstu un paustās emocijas, kā arī to, kā valodas lietojums ietekmē ēdiena uztveri un galu galā – izvēli. Pamatots, ka atkarībā no tā, kā ēdiens tiek nosaukts un kā aprakstītas tā sastāvdaļas un īpašības, tas ietekmē šī ēdiena uztveri (Spence, 2017). Tāpēc ir lietderīgi aplūkot pārtikas aprakstu veidus un analizēt to nozīmi (Kāle et al., 2021). Līdz ar to autore analizēja divas liela mēroga datu kopas: receptes (Kāle & Agbozo, 2021) un sociālā tīkla (*Twitter*) datu kopu latviešu un

angļu valodās (Kāle & Agbozo, 2020a; Kāle & Agbozo, 2020b; Kāle & Rikters, 2021; Kāle et al., 2021).

Jāņem vērā, ka analizējot recepšu datus, nav iespējams iegūt informāciju par pārtikas patēriņa rezultātiem, savukārt sociālo tīklu dati ietver informāciju gan par pārtikas izvēli, gan pārtikas uztveri un tās patēriņu. Receptes ir viens no vienkāršākajiem pieejamajiem apjomīgajiem datu avotiem ar pārtiku saistītu tekstu analīzei, un tādējādi tās ir kalpojušas par vienu no pamatelementiem pārtikas datorikā. Pārtikas datorikas pētījumos receptes galvenokārt ir analizētas, lai salīdzinātu dažādu virtuvju, nacionālo ēdienu un sastāvdaļu saturu, lai noteiktu garšas īpašības. Pamatā ir pieņēmums, ka garša ir svarīga un ka garšu nosaka tieši sastāvdaļu kombinācija (Ahn et al., 2011).

Sociālo tīklu datu analīze, savukārt, ir sasniegusi popularitāti dažādos patērētāju pētījumos (Puerta et al., 2020), kur tiek analizēti valodas dati: *Twitter* gadījumā analīze aptver gan tvītus, gan tēmturus. Sociālo tīklu datu analīze sniedz informāciju par patērētāju pieredzi, izvēloties vai patērējot pārtiku, ļaujot arī izsekot multisensorai pārtikas pieredzei. Svarīgi arī atzīmēt, ka sociālo tīklu analīze sniedz informāciju par to, kā lietotāji sagrupē savas garšas īpašības un savstarpēji saistītās strukturālās saites. Autore uzskata, ka labāka izpratne par valodas lietojumu, kas attiecas gan uz veselīgu, gan neveselīgu pārtiku, var novest pie jaunām metodēm, kā mazināt pretrunas, ka veselīga pārtika ir negaršīgāka par neveselīgo, par ko savā darbā runā Majs un Hofmans (Mai & Hoffmann, 2015). Sīkāks skaidrojums par ar pārtiku saistītās valodas analīzi tvīta tekstos sniegts nākamajā nodaļā kopā ar citām šajā darbā izmantotajām metodēm.

2. PĒTĪJUMA PLĀNS UN METODES

Turpmāk aprakstīts pētījuma plāns un metodes. Nodaļa sākas ar literatūras apskatu un apkopo izvēlēto literatūru no sociālo zinātņu, kognitīvās zinātnes un datorzinātņu jomām. Tālāk analizētas un apspriestas dažādas metodes, kas tika izmantotas autores un līdzautoru veidotajās zinātniskajās publikācijās.

2.1. Literatūras apskats

Turpmākajā nodaļā dots literatūras apskats no sociālo, kognitīvo un datorzinātņu jomām. Tā kā pārtikas tēma ir starpdisciplināra, būtiska bija izpratne par pārtikas patēriņu tieši starpnozaru līmenī. Tālāk seko nozīmīgāko publikāciju saraksts, iekļaujot īsu apskatu par publikācijās paustajām pamatidejām. Ērtības labad literatūras apskats sakārtots attiecīgajās zinātnes disciplīnās: datorzinātņu, kognitīvo un sociālo zinātņu jomās.

Literatūras apskats datorzinātnē

Pārtikas datorika ir starpdisciplināra joma, kas pēta ar pārtiku saistītus jautājumus, izmantojot datorzinātņi. Pārtikas datorikas pārskats, kas tapis kā ACM (*Association for Computing Machinery* - Datortehnikas asociācija) visaptveroša jomas publikācija, sniedz pārskatu par pieejamām datu kopām, datu analīzes metodēm un ilustrē iespējas, ko sniedz pieaugošais liela apjoma datu apjoms pārtikas un datorikas nozarēs (Min et al., 2019a). Kopumā datu kopas ir balstītas vai nu uz attēlu, vai dabiskās valodas datu bāzēm. Tā kā šajā darbā uzmanības centrā ir uz dabiskās valodas balstītas datu kopas, autore tālāk iepazīstinās ar iespējām, ko šādas datu kopas piedāvā labākai izpratnei par pārtikas patēriņu.

Recepšu datu kopa ir bijusi sākotnēji vislabāk pieejamajā un masīvā datu kopa. Viens no nozīmīgākajiem darbiem šajā ziņā bija 2011. gada Āna (*Ahn*) un viņa līdzautoru veiktais pētījums par garšu tīklu un pārtikas savietojamības principiem, kurā norādīts, ka “kulinārijas sastāvdaļu garšas savienojuma (ķīmiskais) sastāvs ir sākumpunkts sistemātiskai sakarību meklēšanai, kas varētu būt sastāvdaļu kombinācijas izvēles pamatā” (Ahn et al., 2011). Kāle un Agbozo (2021) darbs lielā mērā balstās uz šo pieeju, kas noteica, ka Ziemeļamerikas un Rietumeiropas receptēs mēdz savietot sastāvdaļas ar līdzīgiem garšas savienojumiem, bet Austrumāzijas receptēs ir apvienotas sastāvdaļas, kur nepārklājas garšas savienojumi (Jurafsky, 2015). Šajā darbā tika izmantota tēmu modelēšanas metodoloģiju, lai analizētu sastāvdaļu kombinācijas dažādās valstu virtuvēs (Kāle & Agbozo, 2021). Visi šeit minētie pētījumi analizē ēdiena sastāvdaļas no informācijas, kas atrodama receptēs. Šī pieeja sniedz nozīmīgu informāciju par konkrētu valstu un reģionu virtuvju sastāvdaļām, taču no šiem datiem nav iespējams noteikt pašu ēdiena baudīšanas rezultātu – garšas un citu ar ēdiena patēriņu saistītajiem faktoriem. (Ahn et al., 2011)

Sākoties Covid-19 pandēmijai, parādījies arvien lielāks pētījumu skaits par sociālā tīkla teksta datiem. Daudzās valstīs ir analizētas izmaiņas pārtikas patēriņā un paradumos, piemēram, par pamatu izmaiņu noteikšanai izmantojot tvītu jeb mikrobloga ierakstu sociālajos tīklos saturu (Laguna et al., 2020). Lai gan patērētāju pētījumi arvien vairāk fokusējas, lai atšķetinātu un interpretētu tvītu tekstus (Puerta et al., 2020), neviens vēl nav mēģinājis analizēt multisensorās pieredzes tādā veidā, kā tas darīts šī promocijas darba ietvaros. Sīkāk inovatīvā pieeja aprakstīta nākamajā nodaļā, kur apkopotas kognitīvo zinātņu publikācijas.

Literatūras apskats kognitīvajā zinātnē

Šis promocijas darbs lielā mērā balstās uz Spensa (Spence, 2017) darbu, kas uzsver nepieciešamību izprast “visa pārējā” lomu ēdiena patēriņa procesā, mēģinot apzināt, kas ēdienu un dzērienu padara tik baudāmu, stimulējošu un, pats galvenais, neaizmirstamu. Spenss (Spence, 2021) skaidro pārtikas patēriņa sezonālos modeļus, uzsverot, ka gadalaikam un klimatam kopumā ir nozīme tajā, ko patērētāji izvēlas un kā veidojas garšas pieredze. Spenss (Spence, 2015b) arī novērtē skaņas nozīmi pārtikas patēriņā, pētot lomu,

kāda ar ēdiena patēriņu saistītajām skaņām ir mūsu uztverē un multisensoras garšas pieredzes baudīšanā. Kā piemērus šeit var minēt speciāli izveidotu audio-saldinātāju un kartupeļu čipšu kraukšķīguma pastiprinātāju audio formātā. Kraukšķīgums signalizē par pārtikas svaigumu, līdz ar to padarot to kognitīvi patīkamāku. Jāsaprot, ka garša sākas prātā, un vispirms ēdam ar acīm (Spence, 2015a). Tādējādi garšas izpratne sniedzas tālu aiz labi sakombinētu sastāvdaļu saraksta. Tādas multisensoras īpašības kā forma, skaņa, smarža, kā arī konteksts un vide nosaka ēdiena garšu un liek uzdot jautājumus par multisensorās pieredzes būtību un to, kāda ir tās loma pārtikas patēriņā. Autore definē multisensoru pieredzi kā iespaidus, kas satur sensorus elementus, kuru centrā ir konkrēti notikumi un kuri ir atvasināti no tā, kā mēs redzam, dzirdam, sajūtam, izgaršojam un smaržojam pasauli (Velasco & Obrist, 2020). Savukārt, tas, kā nosaukts ēdiens un kā tas aprakstīts, var ietekmēt ēdiena uztveri un tā baudīšanu (Spence, 2017).

Multisensorai pārtikas pieredzei ir izšķiroša nozīme pārtikas uztverē un aprakstošā un afektīvā komunikācijā par to (Kāle et al., 2021). Analizējot vairāk kā desmit gadus apkopotus ēdiena tvītus, iespējams ne tikai izsekot, par kurām multisensorām pieredzēm patērētāji runā vairāk, bet precizēt arī laiku – nedēļas dienas un dienas gaitu (rītu, pusdienas, vakaru). Kā jau norāda kognitīvo zinātņu teorija, patērētāji daudz vairāk spēj izteikties par krāsu nekā, piemēram, par smaržu, jo krāsu aprakstam mums ir daudz konkrētāks valodas leksikons, kuru pielietojam ikdienā. Visas šīs zināšanas var palīdzēt, lai veidotu veiksmīgas sabiedrības veselības komunikācijas kampaņas par pārtikas patēriņa ieteikumiem.

Dalījums starp veselīgu un garšīgu pārtiku mūsu prātos (Mai & Hoffmann, 2015) ir viena no šī darba centrālajām tēmām, jo izpratne par garšīguma un veselīgumu ir kritiski svarīga, lai saprastu faktoros, kas ietekmē pārtikas izvēli. Tā ir viena no centrālajām atziņām arī pārtikas datorikas jomā. Liela daļa ar pārtikas izvēli saistītas lietotnes piedāvā dažādas receptes, atbilstoši indivīda garšas preferencēm. Uz tādiem pašiem pārtikas izvēles principiem balstīti arī pētījumi, kas vērsti uz recepšu sastāvdaļu aizvietojamību, atbilstoši garšas preferencēm. Tātad “garšīguma” aspekts ir nozīmīgs attiecībā uz indivīda izvēlēm gatavot vai patērēt kādu konkrētu ēdienu un zināšanas, ka “garšīgums” konfliktē ar “veselīgumu”, ir nozīmīgas arī, lai analizētu lielos datus attiecībā pret šiem aspektiem un ņemtu vērā to savā ziņā izslēdzošo dabu.

Tomēr jebkurš indivīds, kas veic izvēles, mūsdienu urbānajā un digitālajā vidē mijiedarbojas ar dažādām sociālām grupām, un šīm grupām ir būtiska ietekme uz individuālām izvēlēm, tāpēc nākamajā nodaļā tiks apskatīts teorētiskais pamatojums no sociālo zinātņu perspektīvas.

Literatūras apskats sociālajā zinātnē

Savā darbā par garšu un sociālo klasi Margo Finna (Finn, 2019) apskata sabiedrības garšas veidošanās ceļu sociālo šķiru robežās, analizējot to, kā tas izpaužas caur valodu, kas tiek izmantota ēdiena uztveres raksturošanai. Analizējot sociālo tīklu *Twitter*, pētnieku grupa Abare un līdzautori (Abbar et al., 2015) atklāja dažādas atšķirības attiecībā uz ēdienu patēriņa pieredzēm, ko apraksta pilsētās un ko – laukos. Interesanti, ka “pilsētvidē mēdz pieminēt alkoholiskos dzērienus, bet laukos populāra ir pica un šokolāde” (Abbar et al., 2015), kas kaut kādā mērā signalizē par dzīves stila un statusa jautājumiem dažādās sabiedrības grupās. Lai gan nevienlīdzība un ar to saistītās pārtikas patēriņa atšķirības nav šī promocijas darba pētniecības jautājumu lokā, ir nozīmīgi apzināties nevienlīdzības pastāvēšanu, analizējot liela mēroga sociālo tīklu datus, lai izprastu pētījumu robežas. Šeit viennozīmīgi analizētas pārtikas patēriņa pieredzes tieši salīdzinoši labklājīgā, digitāli prasmīgā sabiedrības grupā, līdz ar to dati nav vispārināmi uz sabiedrību kopumā.

Šāds plašāks pārskats par sociālajiem procesiem nepieciešams, lai izprastu citus ar pārtiku saistītus kritiskus aspektus, piemēram, pārtikas piegādi un loģistiku, kas pēdējo desmitgažu laikā globalizācijas dēļ ir paplašinājusies un saskaras ar tādām problēmām kā garas piegādes ķēdes un pārtikas pusfabrikātu patēriņa pieaugums, kas negatīvi ietekmē veselību, kā tas norādīts stratēģijā “No lauka līdz galdam” (*From Farm to Fork*) (Eiropas Savienības Padome, n.d.). Ielūkojoties pārtikas piegāžu ķēžu vēsturē, redzam, ka pārtikas iepakojuma vēsture seko tās loģistikas vēsturei. Tāpat kā militārie vadītāji centās sinhronizēt armijas kustību ar piegādes ceļiem, tika meklētas arī metodes, kā saglabāt jau piegādātās pārtikas svaigumu. Konservēšana bija pirmā modernā, militāri iedvesmotā iepakojuma tehnoloģija, kas tālāk attīstījās, sākoties industriālajai

revolūcijai astoņpadsmitā gadsimta beigās (Metcalfē, 2019). Tas izskaidro, kāpēc arī mūsdienās pārtikas piegādes ķēdes par prioritāti izvirza pārstrādātu pārtiku ar gariem derīguma termiņiem. Diemžēl tas pasliktina pārtikas patērētāju veselību, tāpēc aizvien lielāks uzsvars ir uz īsu pārtikas piegāžu ķēžu attīstību vismaz ES ietvaros.

Izpratne par šo kontekstu, kurā notiek pārtikas patēriņš, ļauj autorei apspriest sava pētnieciskā darba objektivitātes ierobežojumus, kā arī noteikt pieļaujamo pētniecības rezultātu interpretācijas un vispārināšanas līmeni. Tātad nozīmīgi sociālo zinātņu jautājumi, kādos iespējas apskatīt indivīda un sabiedrības pārtikas patēriņa tendences, ietver labklājības līmeni, dzīvesvietas atšķirības pilsētā vai laukos, kultūras un pārtikas vēstures jautājumus, dažādu sociālo lomu ģimenē aspektus un daudzus citus jautājumus, kas saistībā ar pārtikas izvēli un patēriņu aktualizēti dažādos sociālo zinātņu pētījumos (Finn, 2019; Metcalfē, 2019; Burnett & Ray, 2012).

2.2. Dabiskās valodas datu analīze: datorlingvistika

Receptes, kā viens no visvieglāk pieejamiem lielo datu avotiem ar pārtiku saistītu tekstu agrīnai analīzei, tiek uzskatītas par vienu no pamatelementiem pārtikas datorikā, kas izmanto lielas datu kopas pārtikas patēriņa analīzei. Lielākoties pētnieki koncentrējušies uz dažādu virtuvju, nacionālo ēdienu un sastāvdaļu salīdzināšanu, lai noteiktu garšas īpašības, piemēram, nosakot autentiskas, bieži lietotas sastāvdaļas (Ahn et al., 2011; Min et al., 2019b; Kim & Chung, 2016). Vairumā gadījumu recepšu izpēte bijusi fokusēta uz sastāvdaļu kombināciju un to dažādo reprezentāciju dažādu tautu nacionālajās virtuvēs. Lielu daļu no šiem pētījumiem ir iedvesmojusi lielapjoma digitālo datu pieejamība un iespējas tos izmantot (Kāle & Agbozo, 2021). Autore analizējusi recepšu datu lietderību, izstrādājot jaunu konceptuālu ietvaru ar mērķi piedāvāt jaunus pētnieciskos jautājumus un veidot labāku izpratni par lielajiem datiem, kas saistīti ar pārtiku.

Autore analizējusi *Twitter* ierakstus, kuros pārtikas patērētāji dalījušies ar savām pieredzēm, kas saistītas ar dažādiem ēdieniem un dzērieniem. *Twitter* galvenokārt ir platforma tekstam (nevis attēliem), un to plaši izmanto, lai analizētu dalīšanos ar pārtiku saistītu pieredzi. *Twitter* visvairāk lietotais sociālais tīkls bija no 2011. līdz 2015.gadam, bet pēc tam daudzi lietotāji migrēja uz citiem sociālajiem tīkliem (Kāle et al., 2021). Salīdzinājumā ar citiem datu avotiem, sociālo tīklu dati ļauj mums izsekot spontānām reakcijām, kad cilvēki nekavējoties tvīto, tādējādi izvairoties no iespējamām novirzēm, kas varētu rasties, izmantojot citas viedokļu vākšanas metodes, piemēram, aptaujas (Puerta et al., 2020). Ar “mikrobloga ierakstiem *Twitter*” autore saprot viena tvīta garus/īsus mikrobloga ierakstus. Izpētes vajadzībām tika izmantoti vairāki *Twitter* datu korpusi, tostarp nesen publicētā datu kopa – Latvijas *Twitter* Ēdāju Korpus (LTEC - Sproģis & Rikters, 2020). Dati tika apkopoti desmit gadu laikā, izmantojot 363 ar ēšanu saistītus atslēgvārdus. Datu kopas kopējais lielums pārsniedz 2,3 miljonus tvītu, ko ģenerējuši kopumā 168 703 lietotāji.

Turpmāk aprakstītas pētījumu metodes, kas tika izmantotas sociālo tīklu un citu liela mēroga datu bāzu analīzei. Metodes sagrupētas atbilstoši to mērķim: 1) ar veselīgu pārtiku saistītu hedonistisku konotāciju analīze *Twitter*; 2) ar pārtiku saistītu multisensoro aspektu biežuma noteikšana *Twitter*; 3) ar multisensoro pieredžu saistīto noskaņojumu noteikšana *Twitter*; 4) dažādu nacionālo virtuvju salīdzināšana, izmantojot recepšu datu analīzi.

Ar veselīgu pārtiku saistītu hedonistisku konotāciju analīze *Twitter*

Viens no pieņēmumiem, kas tiek pārbaudīts ar datu analīzes metodēm, ir tas, ka ar veselīgu pārtiku saistītajā diskursā ir maz hedonistisku konotāciju. Respektīvi, par veselīgu pārtiku tiek runāts kā par racionālu, pragmatisku izvēli, taču tā netiek asociēta ar garšīgumu un prieku, to patērējot. Lai atklātu hedonistiskas konotācijas sociālo mediju tīklā *Twitter* saistībā ar veselīgu pārtiku, tika izmantotas datizraces (*data mining*) metodes, kas ir “datu analīzes process, lai aptvertu galvenos jēdzienus un tēmas, kā arī atklātu slēptās attiecības un tendences, iepriekš nezinot precīzus vārdus vai terminus, ko autori izmantojuši šo jēdzienu izteikšanai” (Khatai et al., 2019). Izmantojot *Twitter* API (lietojumprogrammu

programmēšanas interfeisu) R programmēšanas valodā, tika meklētas ziņas angļu valodā par šādām tēmām: #veselīgāpārtika un #dienasrecepte. Lai no datu kopas izvilktu jēgpilnu kontekstuālo informāciju, tika ģenerēti N simboli (Gargouri et al., 2003). Galīgais izmantotais valodas datu analīzes paņēmiens bija vārdu asociāciju noteikšana, lai noskaidrotu tādas asociācijas kā 1) #veselīgāpārtika un atbilstoši atbalsta termini #ilgtspējīga, #barojoša un #garšīga; 2) #vegāns; un #veģetārs, ko atbalsta termini #zemstaukusators, #veselīgs un #garšīgs (Kāle & Agbozo, 2020). Sīkāk iegūtie rezultāti prezentēti pētījuma rezultātu sadaļā.

Ar pārtiku saistītu multisensoro aspektu biežuma noteikšana *Twitter*

Lai sociālajā tīklā *Twitter* noteiktu ar pārtiku saistītu multisensoro aspektu biežumu, tika iegūta, normalizēta un filtrēta liela, domēnam specifiska tvītu datu kopa (ar minēto pārtikas produktu un dzērienu pusautomātiskām anotācijām katram tvītam), lai iegūtu tikai tās vienības, kas tieši attiecas uz izpētes jautājumu. Darbplūsma (*workflow*) tika veidota, balstoties uz šādu uzdevumu veikšanu: 1) identificēt izplatītākos īpašības vārdus un darbības vārdus, ko izmanto, lai aprakstītu dažādu ēdienu (un dzērienu) īpašības, un automātiski iepludināt tos visās iespējamās valodas modifikācijās; 2) izvēlēties tos tvītus, kuros ir vismaz viens no iepriekšminētajiem apzīmējumiem vai darbības vārdiem; 3) atlasīt tikai tos tvītus, kuros minēts arī vismaz viens ēdiens vai dzēriens; 4) izveidot populārāko sakopojumu skaita, laika grafiku un citu attēlojumu analīzi vizuāli. Pēc šo uzdevumu veikšanas, iegūtie rezultāti tika analizēti un interpretēti, nosakot visnozīmīgākās tendences, kas balstītas uz multisensoro aspektu biežumu (Kāle et al., 2021).

Ar multisensoro pieredžu saistīto noskaņojumu noteikšana *Twitter*

Sentimentu, kas saistīti ar pārtikas multisensoro pieredžu aspektiem, noteikšanai sociālo mediju tīklā *Twitter*, autore izmantoja divas dažādas metodoloģijas pieejas. Pirmā pieeja tika izmantota, analizējot tvītus angļu valodā, otrā pieeja – ar pārtiku saistītos tvītus latviešu valodā. Kāles un Agbozo pētījumā (2020b) tika izmantota tekstizrace kā paņēmiens nepazīstamu datu modeļu atklāšanai, izmantojot latentu semantisko analīzi (LSA), lai iegūtu un attēlotu vārdu nozīmi kontekstā, izmantojot statistikas aprēķinus, ko piemēro lielām valodas datu kopām. Tvīti, kuros bija atslēgas vārdi “lapukāposts” un “pārtika”, tika papildus marķēti, lai ar UDPipe modeļa palīdzību noteiktu analīzei nepieciešamos īpašības vārdus (Straka et al., 2016). Tālāk no datu kopas tika iegūtas bigrammas ar īpašības vārdiem. No tīras marķētas datu kopas tika veikta īpašības vārdu bigrammu noskaņojuma analīze. Lai novērtētu vispārīgās emocijas, kas parādās tvītā, pētījumā tika izmantots Lograna-Makdonalda leksikons (*Loughran un McDonald Sentiment Word Lists*), kas sākotnēji tika veidots sentimenta analīzei finanšu analītikas kontekstā (Loughran & McDonald, 2011), taču tika piemērots pārtikas datu analīzei. Otrs izmantotais semantiskais leksikons bija AFFIN semantiskais leksikons, kas katram vārdam piešķir polaritātes punktu skaitu, tādējādi to veido vārdi, kas klasificēti kā pozitīva vai negatīva sentimenta atspoguļojumi.

Atšķirīga pieeja tika izmantota Kāles un Riktera (2021) darbā, kur sentimenta analīzes rīks tika pielāgots un izmantots vispārējās latviešu valodas tvītus (Thakar & Pinnis, 2020). Vienlaikus tika piemērots arī latviešu valodai raksturīgs modelis (Vīksna & Skadiņa, 2020), to iesakņojot visā Latvijas *Twitter* ēdāju korpusā (LTEC). Proti, LTEC tika izmantots, lai noteiktu ar smaržu, garšu un temperatūru saistītus aspektus, izvēloties unikālus tvītus, kuros ir vismaz viens no iepriekš definētajiem īpašības vārdiem vai darbības vārdiem, izmantojot morfoloģijas rīku latviešu valodai (Deksne, 2013). Tas ļāva iemācīt sentimenta klasifikatoru atšķirt negatīvus, pozitīvus un neitrālus tvītus (Kāle & Rikters, 2021).

Dažādu nacionālo virtuvju salīdzināšana, izmantojot recepšu datu analīzi

Lai salīdzinātu dažādu valstu nacionālās virtuves, izmantojot recepšu datu analīzi, tika izmantota tēmu modelēšana. Tā tika uzskatīta par labāko no metodēm, analizējot recepšu datus, jo ļauj pētniekiem noteikt dažādu ēdiena sastāvdaļu nosaukumu izplatību izvēlēto tēmu ietvaros. Tas, savukārt, var vest pie konkrētāku tēmu padziļinātas analīzes (Nikolenko et al., 2017). Darbā tika izmantots par Latento Dirihlē Procesu (LDA) dēvētais tēmu modelēšanas princips, kas, neskatoties uz nespēju modelēt tēmu korelāciju,

ir plaši izplatīts vispārīgs varbūtības modelis. Kā sīkāk skaidrots Kāles un Agbozo (2021) publikācijā, sākotnējā datu kopa, kas iegūta no datu kopas “Yummly-66 k” un sastāv no 66,615 receptēm no desmit virtuvēm vietnē “Yummly” – tīmekļa vietnes, kas ģenerē personalizētas receptes un kalpo kā recepšu meklētājs – ir ēdienu recepšu krātuve. Šī datu kopa tika uzskatīta par piemērotu pētījumam, kurā galvenā uzmanība tika pievērsta Ziemeļamerikas un Meksikas virtuvēm. Tēmu modeļa funkcija tika veidota, pamatojoties uz LDA modeli, lai no korpusa iegūtu desmit (10) tēmu modeļus. Pēc datu kopas tīrīšanas un filtrēšanas, izmantojot runas daļas (POS) marķēšanas modeli (Straka et al. (2016) izstrādāto Udpipeline modeli), Udpipeline modelis tika izmantots, lai atzīmētu POS attiecībā pret visiem marķieriem. Tā tika iegūts rezultāts ar noteiktiem biežākajiem īpašības vārdiem. Visbeidzot, testa datu un apmācības datu kopas tika aplūkotas un rezultāti izanalizēti (Kāle & Agbozo, 2021).

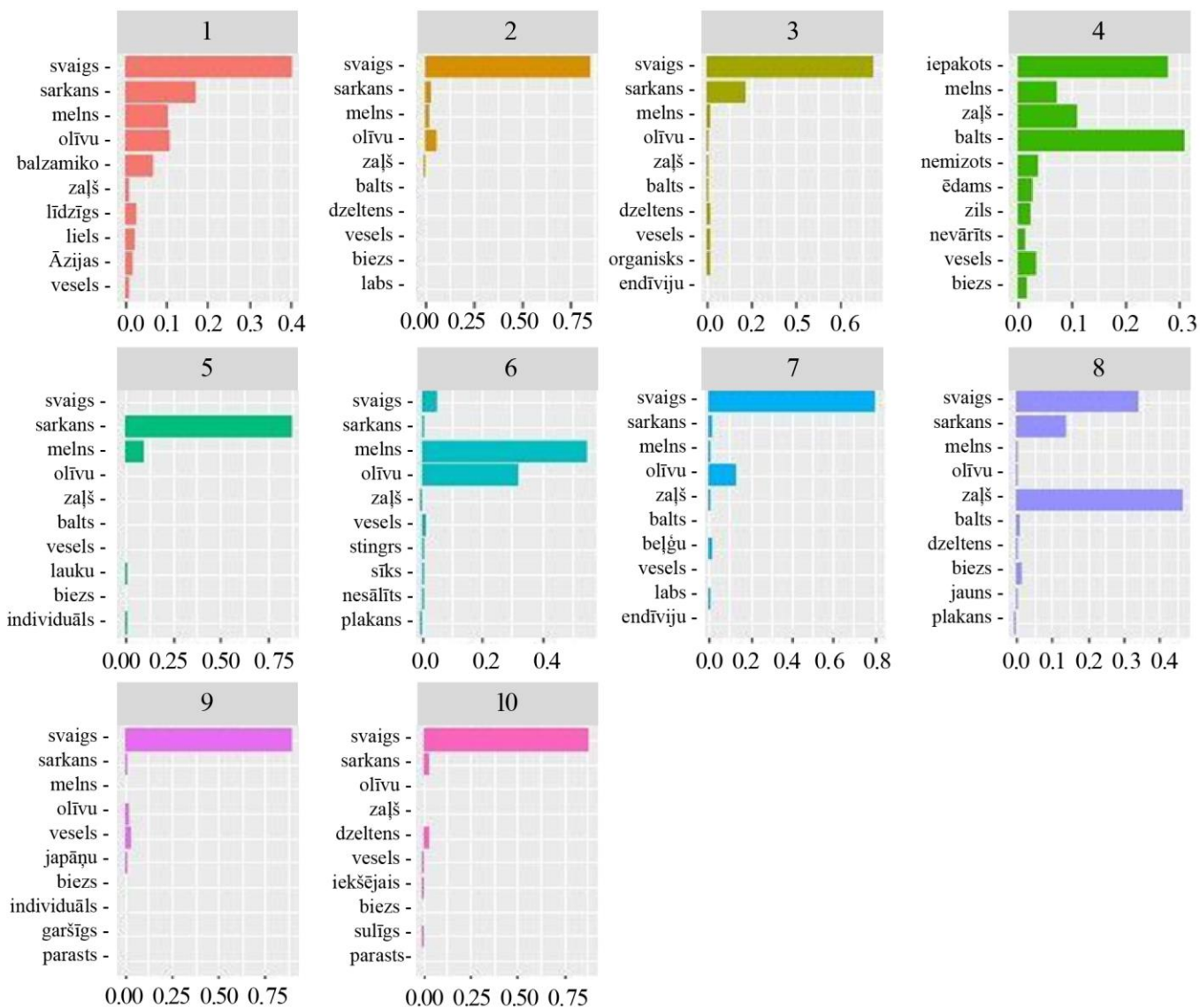
3. KONCEPTUĀLAIS IETVARS

Turpmākajās apakšnodaļās jaunais konceptuālais ietvars (sk. tabulu 1.1.) tiks skaidrots un pamatots, ilustrējot katru no šī ietvara pamatelementiem ar rezultātu piemēru, kas iegūti autores publikāciju kopā ar līdzautoriem, rezultātā. Jau esošie rezultāti tiks papildināti ar nākotnē iespējamu tēmu analīzes apskatu turpmākai pētniecības attīstībai. Konceptuālais ietvars tiks aplūkots secīgi, vispirms izvērtējot izvēlētās tēmas konteksta zināšanu ieguves nozīmību, pēc tam pamatojot nepieciešamību biežuma noteikšanai attiecīgajai tēmai. Kā trešā sadaļa konceptuālajā ietvarā tiks aplūkota emocionāli afektīvo reakciju noteikšana un laika skalas izveide, noslēdzot diskusiju ar nepieciešamību pēc papildu konteksta mainīgo pievienošanas.

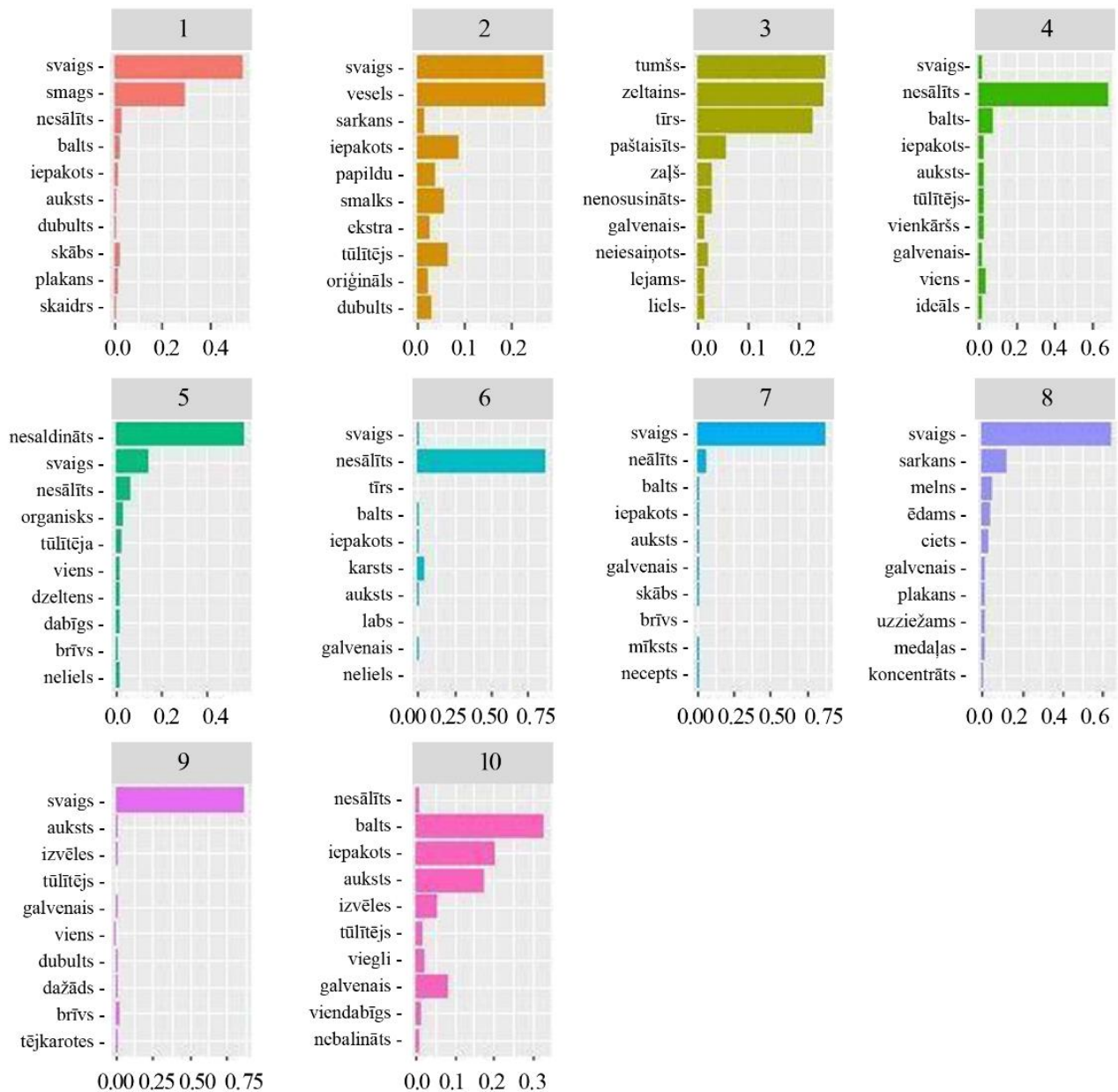
3.1. Izvēlētās tēmas konteksta zināšanu ieguve

Sākotnējais mērķis, salīdzinot dažādas nacionālās virtuves, izmantojot recepšu datus, bija ēdienu klasificēšana veselīgos un neveselīgos pārtikas produktos. Izvirzītais mērķis netika sasniegts, jo veselīgumu pārtikai var noteikt tikai tās lietojuma kontekstā, ņemot vērā gan diētu, gan dzīvesveidu kopumā (Kāle & Agbozo, 2021). Tomēr katrā no nacionālajām virtuvēm, kas tika analizētas, bija iespējams izveidot tēmu modeļus, tādējādi gūstot pārskatu par tur ietvertajiem ēdieniem, to sastāvdaļām un specifiku. Lai gan plaša mēroga recepšu dati var sniegt zināmu pievienoto vērtību pārtikas datorikas pētniekiem, pieņēmumi, ko izmanto, lai analizētu pārtikas patēriņu dažādās kultūrās, tomēr ir pastāvīgi jāpilnveido, pretējā gadījumā iegūto datu analīze nedod pietiekamu pievienoto vērtību (Kāle & Agbozo, 2021).

Padziļinātas konteksta zināšanas nosaka nozīmīgus ierobežojumus, ar ko rēķināties, vispārinot iegūtos datus. Koncentrējoties uz turīgo un digitāli aktīvo sabiedrības daļu, bija nozīmīgi analizēt tendences, kas veido diskusiju tēmas, iekļaujot, piemēram, diskursu par vegānu un veģetāru pārtiku klimata pārmaiņu kontekstā, indivīda veselības un labsajūtas jautājumiem, dažādu nacionālo virtuvju īpatnībām un citiem jautājumiem. Bez kontekstuālas izpratnes par pārtikas sistēmām (Grivins et al., 2020), trūktu nepieciešamā pamatojuma tās izvēlētajam fokusam un padziļinātai ar pārtikas patēriņu saistīto aspektu analīzei.



3.1. att: Īpašības vārdu tēmu modeļi: biežākie īpašības vārdi, kas saistīti ar salātiem Ziemeļamerikas virtuvē (Kāle & Agbozo, 2021)



3.2. att. Īpašības vārdu tēmu modeļi: biežākie īpašības vārdi, kas saistīti ar desertiem Ziemeļamerikas virtuvē (Kāle & Agbozo, 2021)

Šis piemērs no autores publikācijas ilustrē konteksta zināšanu apguvi, atklājot dažādu ēdienu visbiežāk lietotos īpašības vārdus un lietvārdus dažādās nacionālajās virtuvēs. Ilustrācijas nolūkos šeit aplūkojami īpašības vārdu piemēri, kas saistīti ar salātiem un ar desertiem Ziemeļamerikas virtuvēs. Detalizētākas zināšanas par to, kas konkrētajā virtuvē ir visbiežāk lietotie īpašības vārdi un lietvārdi, pētniekiem var palīdzēt precizēt meklēšanas vaicājumus sociālo mediju pētījumiem, kur atslēgas vārdu izvēle ir nozīmīgs pirmais solis pētījumu veikšanā.

Šī specifiskā informācija, kas attiecas uz konkrētu nacionālo virtuvi, var būt noderīga arī salīdzinošos pētījumos, lai noteiktu atšķirības recepšu galvenajās sastāvdaļās, kas veido konkrētu nacionālo virtuvju garšas pamatu. Tēmu modelēšanas metode, kad analizējot dažādus ēdienus (piemēram, salātus, desertus u.c.), iegūst visbiežāk lietoto sastāvdaļu kombināciju pārskatus, pētniekiem ļauj iegūt

ēdiena “metarecepti” konkrētā nacionālajā virtuvē, kā arī veidot virtuvei specifiskus pārtikas produktu un pārtikas sastāvdaļu leksikonu, kas var kalpot par pamatu turpmākai padziļinātai izpētei. Tādējādi tēmu modelēšanas metode lielajiem datiem receptēm un to sastāvdaļām dod nozīmīgus izejas datus, ar kuriem tālāk var strādāt, analizējot, piemēram, sociālo tīklu datus. Receptes un to sastāvdaļas dod ieskatu pārtikas produktos, kas tiek aktīvi izmantoti konkrētā reģionā, un var būt nozīmīga konteksta informācija, analizējot sociālajos tīklos pausto attieksmi pret, piemēram, alternatīvajām olbaltumvielām, gaļas patēriņu vai citiem ar ēdienu saistītiem jautājumiem, kas ir aktuāli dotajā sabiedrībā.

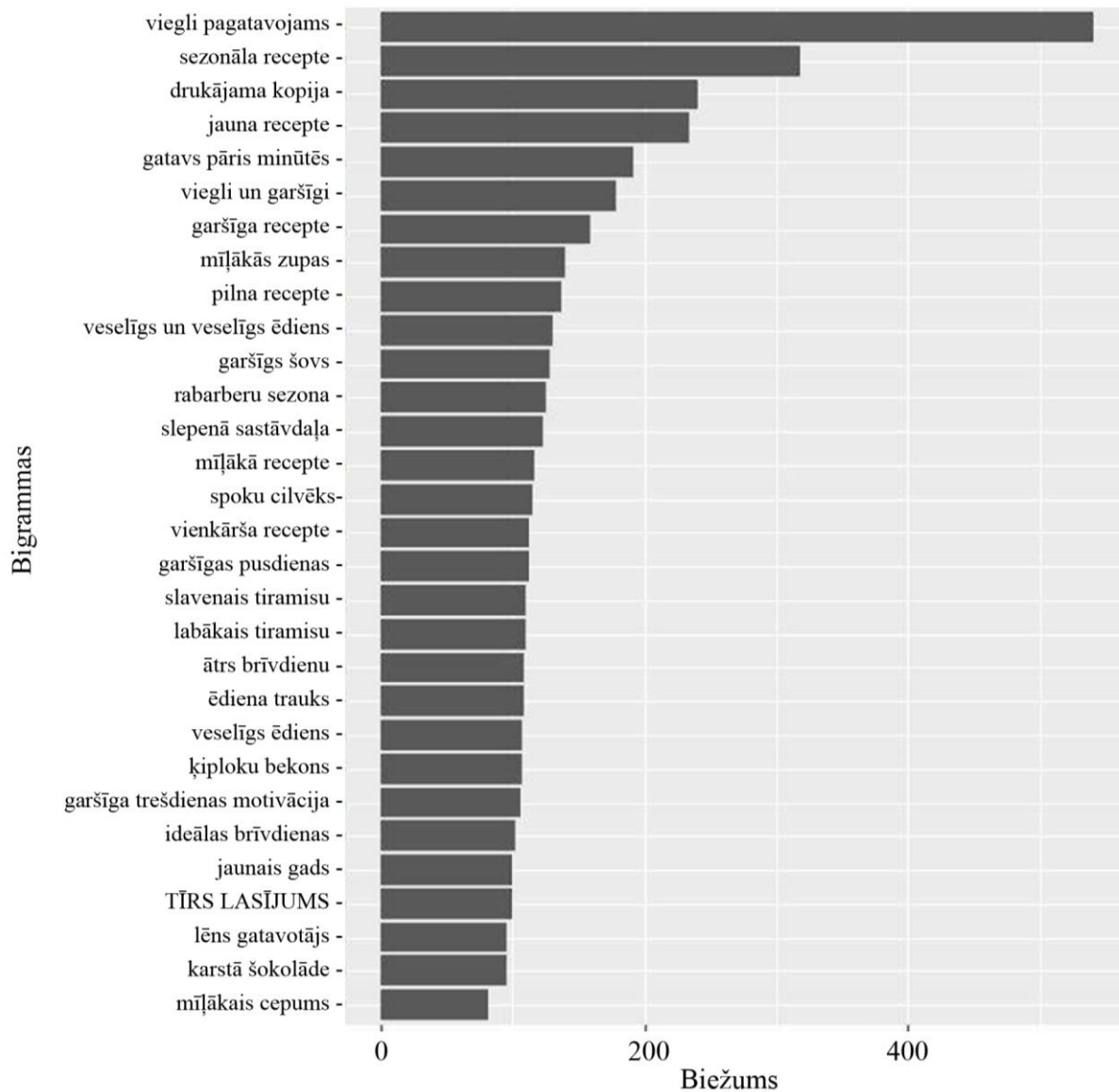
3.2. Biezuma noteikšana

Šī promocijas darba kontekstā īpaša uzmanība tika veltīta “veselīgas” un “garšīgas” pārtikas dalījumam, jo, kā jau izklāstīts literatūras apskata sadaļā, šīs ir viena otru izslēdzošas kategorijas, jo veidojas „neveselīgs = garšīgs intuīcija” (Mai et al., 2011), kas darbojas abos virzienos – veselīga pārtika automātiski tiek uztverta kā negaršīgāka, un neveselīga pārtika kā garšīgāka. Tāpēc kā vienu no galvenajiem pētniecības uzdevumiem šajā sadaļā, autore izvirzīja ar šo tēmu saistīto bigrammu biežuma izvērtēšanu afektīvajām reakcijām attiecībā uz veselīgu pārtiku *Twitter*. Jo biežāk atkārtojas konkrētas bigrammas, jo nozīmīgāks ir diskurss, uz ko šīs bigrammas attiecas. Analizējot lielos datus, biežuma noteikšana ir viena no pirmajām analītiskajām pieejām, ko veikt, lai varētu tos tālāk klasificēt un apstrādāt.

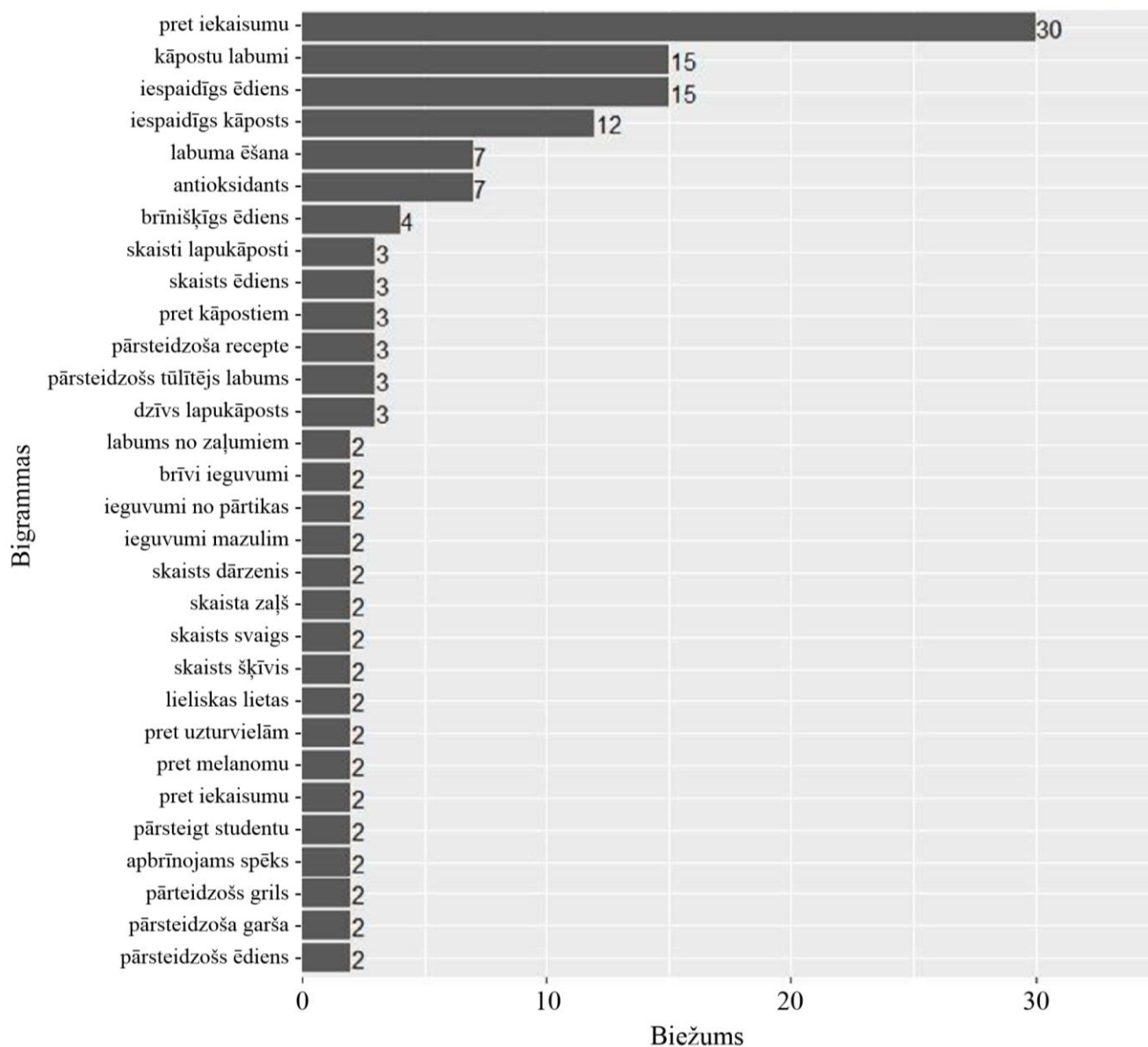
Rezultāti ilustrē, ka vārdu kombinācijai “veselīga pārtika” parasti pievienojas vārdi, kas apzīmē pragmatiskus lēmumus lietot veselīgu pārtiku un ļoti maz tādu vārdu, kas norādītu uz afektīvām reakcijām uz pārtiku (Kāle & Agbozo, 2020a; Kāle & Agbozo, 2020b). Tātad varam secināt, ka veselīgs ēdiens netiek aprakstīts kā garšīgs un baudpilns, bet pamatojums tā patēriņam ir racionālas, pragmatiskas rūpes par savu veselību. Tādējādi pašreizējie tvītu, garo ēdienu blogu ieraksti tvīterī, kas saistīti ar veselīgu pārtiku, nekonzentrējas uz garšas aspektiem un tajos ir maz atsauču uz hedoniskiem izteicieniem, priekšroku dodot “vienkāršai un vieglai” ēdienu gatavošanai, nevis “izsmalcinātai un baudāmai” ēdiena pieredzei (Kāle & Agbozo, 2020a).

Šeit izklāstītās *Twitter* valodas un veselīgas pārtikas sasaiste attiecas arī uz diskusiju par to, cik lielā mērā savietojamas ir valodas datu/tekstu kvantitatīvās analīzes metodes ar pētniecības jautājumiem kognitīvajā zinātnē. Tā vietā, lai pētniecību vērstu tikai uz ēdiena sastāvdaļu analīzi, kā tas līdz šim ir pierasts datorzinātnēs (Ahn et al., 2011), uzmanība būtu jāpaplašina un jāietver, piemēram, ēdiena asociācijas, afektīvas reakcijas uz ēdienu, tāpat arī ēdiena izvēles sasaiste ar sociālo statusu, kā arī citi nozīmīgi konteksta faktori. Paplašinot pētniecības jautājumu loku, varēsīm labāk izprast pārtikas izvēles fenomenu kā tādu (Kāle & Agbozo, 2020.b).

Zemāk piemērā ir attēloti lielāki sastopamības biežumi, kas saistīti ar veselīgu pārtiku *Twitter* angļu valodas tvītos. Kā redzams attēlā 4.3., veselīga pārtika galvenokārt saistīta ar vieglu, sezonālu, ātri pagatavojamu pārtiku raksturojošu bigramu kopumu, kur uzsvērtā racionāla un pragmatiska izvēle, kā arī ieguvumi veselībai. Netiek akcentēti garšas un ēdiena baudīšanas aspekti. Tas savukārt ļauj izdarīt secinājumus, kā sociālajos medijos veidojas tēmas par veselīgu pārtiku un garšīgu pārtiku, un sniedz nepieciešamos pamatelementus, lai saprastu, kā veids, kā mēs runājam par pārtiku, var ietekmēt pārtikas izvēli. Rezultāti, kas uzsver konkrētas ar veselību saistītas īpašības, taču neietver ar labu garšu saistītus aprakstus, redzami arī konkrētu veselīgu pārtikas produktu bigramu analīzē (skatīt attēlu 4.3).



3.3. att. Bigrammu biežums veselīgai pārtikai *Twitter*: sākas ar īpašības vārdu (Kāle & Agbozo, 2020)



3.4. att: Visbiežākās bigrammas attiecībā uz lapu kāpostiem: Bing sentiments un vārds (Kāle & Agbozo, 2020b)

3.3. Emocionāli afektīvo reakciju noteikšana un laika skala

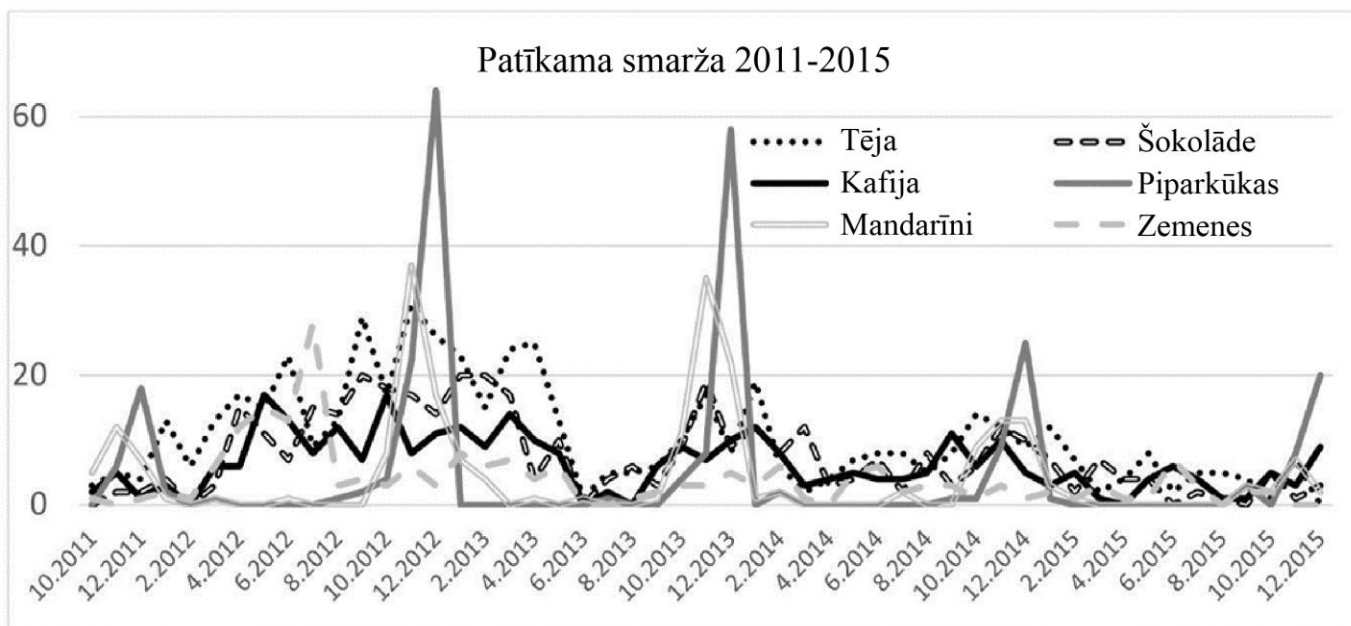
Padziļinātāka emocionāli afektīvo reakciju noteikšana un šo emociju laika skala tika veikta latviešu valodas *Twitter* lietotāju vidū, izmantojot Latvijas *Twitter* ēdāju datu korpusu, novērojot emocionāli afektīvās reakcijas dažādos dienas un nedēļas laikos, izvērtējot multisensoro pieredžu atainojumu. Lai konstatētu afektīvas reakcijas, tika veikta sentimenta analīze (skat. Kāle et al., 2021).

Rezultāti liecināja, ka dažas multisensorās pieredzes ir bieži pieminētas (piemēram, ar krāsu saistītas), bet citas – maz (piemēram, ar tekstūru, smaržu un smaku saistītas). Afektīvu reakciju izvietošana laika skalā ļāva mums secināt par zināmu ar laiku saistītu dinamiku, piemēram, konstatēt, ka kafija un šokolāde arvien vairāk saistās ar “garšīgu” tuvāk nedēļas nogalei, bet pirmdienu rītus ierasts sākt ar veselīgām brokastīm – piemēram, salātiem. No šāda viedokļa Kāle et al., (2021) multisensoro pieredžu pārtikas pētniecībai ir pievienojuši unikālu, tvītu laikspiedola skatupunktu, kas sniedz būtiski informāciju

par to, kā un kad sociālo tīklu lietotāji runā par multisensorām pieredzēm attiecībā uz ēdienu sociālajā tīklā *Twitter*. Šie dati ir nozīmīgi aizvien pieaugošā diennakts ritma analizē attiecībā uz pārtikas patēriņu (Boege et al., 2021). Ņemot vērā indivīdu pārtikas preferences atbilstoši diennakts ritmam, iespējams izstrādāt aizvien precīzākas personalizētas pārtikas lietotnes, ietverot arī grupu dinamiku sociālajos tīklos attiecībā uz ēdiena patēriņa cikliskumu un tā ietekmi uz indivīda pārtikas izvēlēm.

Šis piemērs ilustrē iepriekš aprakstīto Kāle et al., (2021) pētījumu – to, kā konkrēti ēdieni un dzērieni tiek aprakstīti to multisensoro īpašību kontekstā – šajā piemērā ar patīkamas smaržas afektīvām reakcijām. Virsotnes signalizē pārtikas lietojuma biežuma pieaugumu atkarībā no sezonālā ritma. Virsotņu augstumu samazinājumi skaidroti ar to, ka *Twitter* lietotāji kļūst mazāk aktīvi un to skaits samazinās, jo daudzi 2015. gadā migrē uz *Instagram* un citiem sociālajiem tīkliem. Analizējot pārtikas un multisensoro pieredžu reprezentācijas sastopamības biežumu, izdevies atklāt sezonālā ritma, kā arī apzināt atsevišķus kultūras aspektus, kas saistīti ar tradīcijām konkrētā nacionālajā virtuvē, šajā gadījumā – latviešu. Tā ir arī atklājusi ar pārtiku saistītu tvīta darbību attiecībā uz konkrētām nedēļas dienām un konkrētiem dienas laikiem, atklājot tvīta darbību attiecībā uz konkrētiem pārtikas produktiem konkrētos dienas laikos (Kāle et al., 2021).

Zināšanas un augstāks precizitātes līmenis par konkrētiem pārtikas produktiem, kas iezīmējas sezonālā ritmā, var palīdzēt sabiedrības veselības politikas veidotājiem veidot tādas komunikācijas stratēģijas, kas rezonēs sociālajos medijos. Savukārt, datorzinātnēs sezonālā ritmam un cikliskumam attiecībā uz dažādu pārtikas produktu patēriņu ir aizvien lielāka nozīme, izstrādājot personalizētas pārtikas ieteikuma lietotnes un veidojot ieteikumu sistēmas (Rosa et al., 2022).



3.5 att.: Patīkamas smaržas sentimenta grafiks izvēlētiem pārtikas produktiem un dzērieniem, 2011.–2015. gads (Kāle et al., 2021)

	Čāpstināt	Kraukšķēt	Sprakšķēt	Šņakstināt	Čurkstēt
Čipsi	5	15	0	0	0
Šokolāde	0	11	4	0	0
Gaļa	2	7	3	0	3
Kartupeļi	0	11	0	0	2
Ābols	4	8	0	0	0
Piparkūkas	1	8	3	0	0
Krēmzupa	0	11	0	0	0
Bekons	0	11	0	0	0
Tēja	3	4	3	0	0
Salāti	1	7	0	1	0
Mērce	0	9	0	0	0
Cūkgaļa	3	5	0	0	1
Saldējums	0	4	2	0	2
Zupa	0	8	0	0	0
Siers	0	8	0	0	0
Vista	2	2	1	0	2
Ūdens	1	4	1	0	1
Ķirbis	0	6	0	1	0
Pankūka	3	2	0	0	0
Zemenes	0	2	2	0	1

3.6 att.: Top 20 ēdieni un dzērieni, kas sakrīt ar pieciem visbiežāk lietotajiem darbības vārdiem, kas apraksta skaņas (Kāle et al., 2021)

3.4. Papildu konteksta mainīgo pievienošana

Pēdējais solis jaunajā konceptuālajā pārtikas patērētāju izpratnes ietvarā ir jaunu konteksta mainīgo lielumu pievienošana, kas atklāti, veicot iepriekšējo analīzi. Autore pievieno divus jaunus konteksta mainīgos lielumus, ko atklājusi iepriekšējā analīze: laika apstākļus un veselīguma uztveri. Tie sīkāk aprakstīti zemāk.

3.4.1. Laikapstākļu dati

Digitālie sociālie tīkli ir vērtīgi, jo ir jutīgi pret sabiedrības tendencēm un nodrošina jaunus veidus, kā izsekot sezonālām garšu un pārtikas produktu atšķirībām, kas īpaši redzamas valstīs, kurās temperatūras ziņā ir lielākas temperatūras svārstību klimats (Spence, 2021). Latvija ar saviem četriem dažādiem gadalaikiem šajā gadījumā nodrošina labu piemēru tvītu informācijas analizēšanai aukstajā ziemas sezonā un arvien karstākajās vasaras dienās (Kāle et al., 2021).

Sasaistot ar pārtiku saistītos datus no *Twitter* ar laika apstākļu datiem, būtu iespējams izvērtēt saistību starp pārtikas izvēli un ar pārtiku saistīto valodu, no vienas puses, un laika apstākļiem un to publisko vērtējumu/uztveri, no otras puses (Bakhshi et al., 2014; Bujisic et al., 2019).

Šīs zināšanas par tvītu biežumu un iegūto noskaņojumu var būt noderīgas sabiedrības veselības politikas veidotājiem, un tās var izmantot, mudinot patērētājus izvēlēties veselīgākas pārtikas alternatīvas dažādos laika apstākļos un gadalaikos. Laikapstākļu ietekmes uz pārtikas patērētājiem un viņu afektīvo reakciju atzīšana un izpratne palīdz izskaidrot arī dažādas neefektivitātes, kas saistīti ar pārtikas patēriņu – pārtikas izšķērdēšanu, veselīgas vai neveselīgas pārtikas izvēli un citus jautājumus.

Šajā publikācijā radītās kontekstuālās zināšanas var būt noderīgas datorzinātņu pētniekiem, kuri strādā ar personalizētu pārtikas un veselības lietotņu izstrādi, jo cilvēki ir sociālas būtnes un apkārtējo uzvedība ietekmē viņu izvēli. Laikapstākļu datu sasaiste ar pārtikas izvēlēm būtu noderīga, veidojot Pasaules pārtikas atlantu, kas ir atvērta pirmkoda platforma, kas satur informāciju par pārtikas pieejamību, ņemot vērā dažādus faktoru ar mērķi ilustrēt to, kā pārtikas un dzīvesveida navigācijas sistēmas var izmantot pasaules pārtikas atlantu, lai nodrošinātu personalizētus un uz kontekstu orientētus risinājumus (Rostami et al., 2022).

Turklāt šis pētījums var palīdzēt izcelt mūsu digitālās un analogās dzīves savstarpējās saiknes līmeni — sekot līdzi tvīta noskaņojumam un biežuma rādītājiem, izmantojot faktiskos pirkšanas paradumus un pārtikas pārdošanas datus.

Lai varētu veikt pētniecību par laikapstākļu ietekmi uz pārtikas izvēlēm, nepieciešami pieejami detalizēti laikapstākļu dati, lai varētu veikt statistiski nozīmīgu analīzi attiecībā uz pārtikas tvītu sentimenta analīzi.

3.4.2. Veselīguma uztveres noteikšana ar datorlingvistikas metodēm

Kā viens no svarīgākajiem turpmākiem soļiem ar pārtiku saistītu plaša mēroga datu analīzei ir veselīguma uztvere. Dažādu pārtikas veidu klasificēšana “veselīgos” un “neveselīgos” tomēr ir izaicinoša. Veselīgums kā jēdziens ir atkarīgs no daudziem aspektiem, kas ir daudz plašāki nekā tikai ar pārtiku saistīti faktori. Vispārējs novērojums ir tāds, ka “veselības un garšas mērķi bieži vien konfliktē, un garša parasti dominē pārtikas izvēlē” (Mai et al., 2011). Vārdiem, kas raksturo ēdienu, ir izšķiroša nozīme šī ēdiena pozicionēšanā kognitīvajā kartē, un, raksturojot ēdienu kā veselīgu, samazinās priekšstats par tā garšu “neveselīgs = garšīgs” intuīcija (Mai & Hoffmann, 2015). Tāpēc intriģējošs virziens turpmākajiem pētījumiem varētu būt saistīts ar veselīgu pārtiku vai to, ko cilvēki parasti raksturo kā veselīgu pārtiku attiecībā uz savu veselību vai vides ilgtspēju (Kāle et al., 2021).

Pārtikas izvēlei un pārtikas patēriņam ir liela nozīme sabiedrības veselībā, kā arī tie būtiski ietekmē vides ilgtspējību. Lai gan pārtikas ietekme uz indivīda veselību ir joma, kam pārtikas politikas veidotāji un uztura speciālisti pievērsušies visā pasaulē, ir parādījies jauns diskurss saistībā ar pārtikas patēriņu, proti, ietekme uz planētas veselību vai bioloģiskās daudzveidības, piesārņojuma un CO₂ līmeni, kas ietekmē un veido klimata pārmaiņas, kā arī planētas ekosistēmas kopumā (Grivins et al., 2020). Trešdaļa no globālajām oglekļa dioksīda emisijām tiek attiecināta uz pārtikas sistēmām, kur lielāko devumu dod lauksaimniecība un zemes izmantošana (aptuveni 71% no kopējām emisijām), savukārt pārtikas piegādes ķēde — transports, patēriņš, mazumtirdzniecība un citi saistīti procesi — veido attiecīgi 29% (Crippa et al., 2021).

Jaunās uztura politikas ietver ne tikai ieteikmi uz indivīda veselību, bet apskata arī pārtikas patēriņa ieteikmi uz vidi un planētas ilgtspēju. 2019. gadā EAT-Lancet komisija, ko vadīja 37 vadošie zinātnieki visā pasaulē, izveidoja pasaulē pirmās veselīga un ilgtspējīga uztura vadlīnijas, kas pazīstamas kā Planetārās veselības diēta (PHD), kas iesaka vairāk nekā divas reizes palielināt augļu, dārzeņu, pākšaugu un riekstu patēriņu, kā arī par vairāk nekā 50% samazināt pievienotā cukura un sarkanās gaļas daudzumu (Willet et al., 2019). Gaļas ražošanai ir vislielākā ietekme uz siltumnīcefekta gāzu ražošanu, jo tā veido gandrīz 60% no visām siltumnīcefekta gāzēm, kas rodas pārtikas ražošanā. Liellopu gaļa rada vienu ceturtdaļu no kopējām emisijām, un kopumā dzīvnieku izmantošana gaļas ieguvei rada divreiz lielāku piesārņojumu nekā augu izcelsmes pārtikas ražošana (Xu et al., 2021).

Šeit sociālo mediju datu analīze var būt instrumentāla, lai noteiktu sabiedrības noskaņojumu attiecībā uz ilgtspējīgas pārtikas patēriņa paradumiem, piemēram, analizējot ar gaļas patēriņu saistītus tvītus, papildus analīzi veicot par tvītiem, kas attiecas uz alternatīvajām olbaltumvielām, veģetāru un/vai vegānu

dzīvesveidu un rūpēm ar klimata krīzi, vides ilgtspēju un bioloģisko daudzveidību. Arī šeit svarīgi ņemt vērā klimata apstākļus, kuros konkrētā sabiedrība atrodas, tāpat arī kultūras un nacionālo virtuvju īpatnības. Veselīguma uztveres noteikšana attiecībā uz indivīdu un vidi ir nozīmīga arī izstrādājot pasaules pārtikas atlantu, kur šādas konteksta zināšanas ir būtiskas, lai veidotu visaptverošu lielo datu sistēmu attiecībā uz pārtiku (Rostami et al., 2022).

4. IEROBEŽOJUMI

Viens no ierobežojošiem faktoriem, lai iegūtu padziļinātu izpratni par multisensoriem pārtikas aspektiem, kā attēlots sociālajos medijos, ir leksikona par pārtiku trūkums, ko datorzinātnieki varētu izmantot, nosakot atlasas kritērijus sociālo tīklu teksta datos. Šādus leksikonus būtu jāattīsta atbilstoši sociālo tīklu valodai un kontekst-jūtīgi (piemēram, ņemot vērā pilsētvidi un reģionu specifiku pārtikas kultūrā). Leksikoni, kas tika izmantoti, bija atvasināti no angļu valodas leksikoniem, kas nozīmē to, ka tulkojumā varētu būt pazuduši daži svarīgi aspekti.

Ar pārtikas terminiem saistītu leksikonu izstrāde iet roku rokā ar liela mēroga datu digitalizāciju dažādās vietējās valodās, kas nozīmīgi paplašinātu liela mēroga datu analīzes iespējas. Ja būtu bijusi piekļuve digitāliem recepšu datiem latviešu valodā, tad tēmu modelēšanas analīzei mēs būtu varējuši to izmantot, taču šādu datu trūkuma dēļ, tika veikta Ziemeļamerikas un Meksikas virtuvju tēmu modeļu analīze.

Tā kā pārtikas patēriņš ir sezonāli saistīts un cieši saistīts ar valsts pārtikas kultūru, pārtikas datorikas kontekstā kopumā būtu nepieciešama padziļināta sadarbība ar pārtikas antropoloģijas ekspertiem, kuri varētu sniegt zināšanas sociālo mediju datu meklēšanas vaicājumiem, kas saistītas ar tradīcijām, uzskatiem un niansētāku sabiedrības pārtikas patēriņa ainu. Paši recepšu dati gan nozīmē ierobežotu iespēju analizēt patiesos ēdiena gatavošanas paradumus, jo receptes ir stingri sadalītas nacionālo virtuvju ietvaros, taču mūsdienās šīm lietām globalizācijas ietekmē ir tendence sajaukties, un labākus rādītājus par ēšanas paradumiem varētu nodrošināt pilsētu un lauku reģionu pārtikas patēriņa paradumu analīze (Grivins et al., 2020).

Vēl viens nozīmīgi ierobežojošs faktors ir pievēršanās turīgam, digitāli literātam superpatērētājam, kas ir aktīvs/a sociālo mediju tīklos, nevis sabiedrībai kopumā. Šāda zināšanu sadrumstalotība ir raksturīga lielajiem datiem. Tomēr neraugoties uz sadrumstalotību, liela mēroga sociālo tīklu dati joprojām var sniegt nozīmīgas zināšanas par digitāli aktīvu sabiedrības grupu uzvedību. Fragmentētības dēļ izvēloties neanalizēt plaša mēroga sociālo tīklu datus, pētnieki nevar attīstīt jaunas metodoloģijas un pieejas, lai trūkumus novērstu un tomēr izmantotu vērtīgos, lai arī fragmentētos datus. Autore neanalizēja, kā digitālā uzvedība ietekmē analogo dzīvi, taču šāds pētījums būtu likumsakarīgs turpinājums iepriekš veiktajam.

Visbeidzot, nozīmīgas datu kopas (piemēram, patērētāju pirkšanas dati, precīzi laikapstākļu dati, ar veselību saistīti dati) nav pieejamas, lai varētu izveidot visaptverošu ar pārtikas patēriņu saistītu analīzi.

NOSLĒGUMS

Promocijas darba izstrādes rezultātā tika izveidots **jauns konceptuāls ietvars** pārtikas uztveres un patēriņa analīzei, izmantojot lielo datu sniegtās iespējas. Darba rezultāti nozīmīgi papildina pārtikas datorikas attīstību, piedāvājot dažādas metodoloģiskas pieejas dabiskās valodas lielo datu analīzei par pārtiku, tās uztveri un patēriņa tendencēm. Jaunais konceptuālais ietvars sastāv no četriem galvenajiem blokiem:

1. konteksta zināšanu iegūšana par tēmu;
2. tēmas biežuma konstatēšana;
3. afektīvu reakciju noteikšana un laika grafika analīze;
4. papildu konteksta mainīgo pievienošana.

Konteksta zināšanu iegūšanai nepieciešams veidot globālu pētnieku tīklu, kas pārstāv dažādas pētniecības jomas attiecībā uz pārtiku

Te uzsvērta nepieciešamība koncentrēties uz konteksta zināšanu, kas ietekmē pārtikas sistēmas un sabiedrību, iegūšanu, tādējādi attīstot svarīgus pētniecības virzienus, kas iepriekš nav bijuši pētnieku redzeslokā. Pārtikas kultūras dažādās valstīs un reģionos atšķiras, un strādājot ar lielajiem datiem par pārtiku, tas jāņem vērā. Pārtikas dati, kas iegūti vienā jomā, reti ir derīgi citā (Yamakata et al., 2022). Līdz ar to, veicot pētniecību, ir būtiski iekļaut analīzi par valstu vai reģionu pārtikas sistēmām un izaicinājumiem, ar kuriem tās sastopas. Tāpat svarīgi izprast sociālekonomiskos datus un sabiedrības digitālo prātību, jo sociālo mediju datus rada tikai salīdzinoši labklājīga un digitāli prasmīga sabiedrības daļa – tātad šie ir ierobežojumi, ar kuriem jāreķinās, analizējot lielos datus par pārtikas izvēlēm un patēriņu. Pārtikas datorikas turpmākajā attīstībā būtiski iekļaut šos ar konteksta zināšanām saistītos faktoros un viens no veidiem, kā to varētu nodrošināt, ir starpdisciplināru pētniecības komandu veidošana, kur ar datorzinātniekiem sadarbotos sociālo zinātņu virziena pētnieki – sociologi, pārtikas sistēmu vēsturnieki un antropologi. Līdzšinējā datorzinātņu pētnieku sadarbība ar uzturzinātņu speciālistiem ir atzīstama, taču ne pietiekama, lai apzinātos būtiskos konteksta faktoros, kas ir nozīmīgi, analizējot lielos datus par pārtiku dažādās sabiedrībās. Līdz ar to ir nepieciešams veidot globālu pārtikas datorikas pētnieku tīklu, kas iekļautu pētniekus no dažādām zinātņu disciplīnām.

Tēmas biežuma konstatēšanai sociālo mediju tīkli būs nozīmīgākais avots arī turpmākā pārtikas datorikas pētniecībā, neskatoties uz šo sociālo mediju tīklu fragmentētību un daļējam iespējam rezultātus vispārināt

Pārtika ir viena no galvenajām tēmām, kas tiek apspriesta sociālajā tīklā *Twitter*. Kā platforma galvenokārt tekstam, nevis attēliem, un pateicoties tā pieejamībai pētniecības nolūkos, *Twitter* ir kļuvis iespējams sekot līdz visnējaušākajām tvītotāju ikdienas dzīves detaļām, tostarp informācijai par to, kas, kā un kur viņi ēd (Kāle et al., 2021). Salīdzinājumā ar citiem analīzes avotiem, sociālo mediju dati ļauj mums izsekot spontānām reakcijām, kad cilvēki bez kavēšanās raksta *Twitter*, tādējādi izvairoties no iespējamām aizspriedumiem, kas varētu būt radušies, izmantojot citas viedokļu vākšanas metodes, piemēram, aptaujas un pārtikas dienasgrāmatas (Laguna et al., 2020). Piedevām, sociālo mediju dati ir lielapjoma dati, kas ļauj konstatēt biežāk pārrunātās tēmas un jautājumus, kas saistīti ar pārtiku.

Tēmu biežuma konstatēšana ļauj izprast, kas dominē diskusijās par dažādu pārtiku. Viens no vērtīgākajiem dalījumiem, kā apskatīt dabiskās valodas datu rezultātus attiecībā uz pārtiku, ir dalījums veselīgā un neveselīgā pārtikā. Zināšanas par to, kā cilvēki apraksta veselīgu un neveselīgu pārtiku no datorzinātņu perspektīvas ir nozīmīgas dažādu personalizēto pārtikas ieteikuma lietotņu izstrādē, kā arī attīstot Pasaules pārtikas atlantu (Rostami et al., 2022). Ņemot vērā kognitīvo zinātņu pētījumus, kas norāda uz to, cik lielā mērā vārdi ietekmē uztveri par pārtiku, datorlingvistikas speciālistiem paveras plašs pētniecības lauks izvērst jaunu pētniecības uzdevumu izstrādi, balstoties uz zināšanām par dabiskās valodas datiem attiecībā uz pārtiku.

Afektīvu reakciju noteikšanā un laika grafika analīzē noteicoša loma būs lielo datu kvalitātei

un pieejamībai

Te ilustrēti piemēri par to, kā apvienot kognitīvās zinātnes atklājumus, piemēram, izpratni par afektīvām reakcijām uz pārtiku un multisensoro pieredzi, kas saistīta ar pārtikas patēriņu, ar datorzinātnē izmantotajām metodēm, lai analizētu liela mēroga datus. Liela daļa jaunāko lietotņu pārtikas un veselības jomā balstītas uz diennakts ritmu attiecībā uz pārtikas patēriņu, līdz ar to labāka izpratne par cikliskumu un ar pārtiku saistīto emociju atainojumu, ir nozīmīga. Tomēr ierobežojošs faktors ir datu pieejamība. Kamēr nav pietiekami lielas un detalizētas datu kopas, precīza analīze nav iespējama. Tā, piemēram, lai korelētu lielo datu par pārtiku sentimenta analīzi ar laikapstākļu datiem, nepieciešami detalizēti laikapstākļu dati – gan diennakts laika mērījumu, gan ģeogrāfiskā reģiona zināšanu kontekstā. Līdzīgi ar datiem, kas saistīti ar indivīda veselību. Kamēr pārtikas datorikas pētniekiem ir daļēja piekļuve nozīmīgām datu bāzēm, kas saistītas ar cilvēka veselību un dzīves stilu, tikmēr pētījumu rezultāti būs aptuveni. Līdz ar to viens no pirmajiem pārtikas datorikas pētnieku tīkla uzdevumiem būtu veidot un uzturēt atvērto datu kopas, kuras varētu izmantot pārtikas datorikas analīzes uzdevumu veikšanai.

Papildu konteksta mainīgo pievienošana būs iteratīvs process, ņemot vērā konkrētu reģionu sociālpolitiskos un klimata faktorus

Starpdisciplinārā pieeja uzlabo izpratni par pārtikas patēriņu un sniedz papildu pētniecības impulsu pārtikas datorikas pētniekiem. Lai arī līdz šim pārtikas datorika galvenokārt bijusi vērsta uz ēdienu sastāvdaļu analīzi, to savietojamību un uzturvērtību (Min et al., 2019a, Ahn et al., 2011; Kim un Chung, 2016), nozīmīgu pavērsieni veidojuši Jurafsky (2015) un Fenko et al. (2010) darbi, kas veicinājuši valodas analīzes saistībā ar pārtikas uztveri attīstību. Attiecībā uz pārtikas datoriku un valodas datu analīzi tomēr pastāv starpdisciplināritātes plaisa, jo nozīmīgi kognitīvo zinātņu jautājumi nav pētīti šo zinātnes nozaru metodoloģiju ietvaros. Līdz ar to, ir būtiski ņemt vērā analizējamo lielo datu ģeogrāfijas un klimata jautājumus, kā arī iekļaut reģionā nozīmīgu politikas dokumentus, kam var būt ietekme uz reģiona pārtikas sistēmu attīstību.

Jaunais konceptuālais ietvars papildina pārtikas datorikas pētniecības metodoloģiju un sniedz piemēru jaunu pieeju pētniecības jautājumu izstrādē

Šajā promocijas darbā izvirzītais galvenais uzdevums bija izstrādāt jaunu konceptuālu pētniecības ietvaru, kas veidots, balstoties uz autores un līdzautoru publikācijām, kas risina nozīmīgus kognitīvo un datorzinātņu pētniecības izaicinājumus, lai labāk saprastu lielos datus par pārtikas patēriņu. Autores pētnieciskais darbs ilustrē to, ko ir iespējams sasniegt, pārtikas patēriņam piemērojot starpnozaru domāšanu. Uzdod šādu izpētes jautājumu: **kādas metodes, izmantojot liela mēroga datus, pielietot, lai labāk izprastu pārtikas patērētājus?** Autore motivēja plaša mēroga sociālo tīklu datu pieejamība un nepieciešamība apvienot starpdisciplināras zinātnes jomas, kas var izskaidrot pārtikas patēriņu un veicināt veselīgāku pārtikas patēriņu un līdz ar to arī veselīgāku sabiedrību. Autore ir atrisinājusi dažas pētniecības nepilnības un sniegusi skaidru metodoloģisku sistēmu, kā turpināt darbu ar starpdisciplināriem pētījumiem, vienlaikus izmantojot pievienoto vērtību, ko liela mēroga dati var sniegt šādai analīzei.

ATSAUCES

Interneta resursi

- (2022). *CEA++ '22: Proceedings of the 1st International Workshop on Multimedia for Cooking, Eating, and related APPLications*. New York: Association for Computing Machinery.
- ACM Digital Library. (n.d.). *CHI: Conference on Human Factors in Computing Systems*
<https://dl.acm.org/conference/chi>
- Allmendinger, J. (2015). *Quests for interdisciplinarity: A challenge for the ERA and HORIZON 2020*.
<https://commission.europa.eu/system/files/2019-04/allmendinger-interdisciplinarity.pdf>
- Eiropas Parlaments. (2021). *Lielie dati: definīcija, priekšrocības, sarežģījumi (infografika)*:
<https://www.europarl.europa.eu/news/lv/headlines/priorities/maksligais-intelekt-es/20210211STO97614/lielie-dati-definicija-prieksrocibas-sarezgijumi>
- Eiropas Savienības Padome. (n.d.). *No lauka līdz galdam*:
<https://www.consilium.europa.eu/lv/policies/from-farm-to-fork/>
- Personalized nutrition startup Zoe closes out Series B at \$53M total raise* (2021). TechCrunch:
[shorturl.at/AXY06](https://www.techcrunch.com/2021/09/22/zoe-nutrition-series-b/)
- The PREDICT program. (2020, October 1). JoinZoe: <https://joinzoe.com/whitepapers/the-predict-program>
- UC Santa Cruz. (n.d.). *Making Sense of Human Food Interaction*:
<https://www2.ucsc.edu/hfi/conceptualmodel.html>
- World Economic Forum. (2020, September 20). *How digital technology can accelerate food sustainability*: <https://www.weforum.org/agenda/2022/09/how-digital-technology-can-accelerate-food-sustainability/>
- World Health Organization. (2021, October 20). *Making diets environmentally friendly: Nordic countries lead the way*: <https://www.who.int/europe/news/item/20-10-2021-making-diets-environmentally-friendly-nordic-countries-lead-the-way>

Literatūras saraksts

- Abbar, S., Mejova, Y., & Weber, I. (2015). You tweet what you eat: Studying food consumption through twitter. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, pp. 3197–3206, New York, NY, USA. Association for Computing Machinery.
- Ahn, Y.-Y. et al. (2011). Flavor network and the principles of food pairing. *Scientific Reports*, 1(1), p. 196. doi:10.1038/srep00196.
- Andersen, T., Byrne, D. V., & Wang, Q. J. (2021). How Digital Food Affects Our Analog Lives: The Impact of Food Photography on Healthy Eating Behavior. *Frontiers in psychology*, 12, 634261.
- Bakhshi, S., Kanuparth, P., & Gilbert, E. (2014). Demographics, weather and online reviews. *Proceedings of the 23rd International Conference on World Wide Web*. <https://doi.org/10.1145/2566486.2568021>
- Boege, H. L., Bhatti, M. Z., & St-Onge, M.-P. (2021). Circadian rhythms and meal timing: impact on energy balance and body weight. *Current Opinion in Biotechnology*, 70:1–6. Food Biotechnology Plant Biotechnology.
- Bujisic, M., Bogicevic, V., Parsa, H. G., Jovanovic, V., & Sukhu, A. (2019). It's raining complaints! How weather factors drive consumer comments and word-of-mouth. *Journal of Hospitality & Tourism Research*, 43(5), pp. 656–681.
- Burnett, S. C. & Ray, K. (2012). Sociology of Food. *The Oxford Handbook of Food History*. Oxford University Press.
- Carolan, M. (2015). Publicising food: Big data, precision agriculture, and co-experimental techniques of addition. *Sociologia Ruralis*, 57:135–154.
- Crippa, M., Solazzo, E., Guizzardi, D., Monforti-Ferrario, F., Tubiello, F.N., & Leip, A. (2021). Food systems are responsible for a third of global anthropogenic GHG emissions. *Nature Food*.
- Deksne, D. (2013). Finite State Morphology Tool for Latvian. *Proceedings of the 11th International Conference on Finite State Methods and Natural Language Processing*. St Andrews, Scotland: Association for Computational Linguistics, pp. 49–53. Available at: <https://aclanthology.org/W13-1809> (Accessed: 13 September 2021). doi:10.3389/fcomp.2019.00003.
- Fenko, A., Otten, J., & Schifferstein, H. (2010). Describing product experience in different languages: The role of sensory modalities. *Journal of Pragmatics*, 42(12), pp. 3314–3327.
- Ferran B., Jhaveri S., Lutz R., Isbister K. & Wilde D. (2019). CHI '19: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–13. doi: <https://doi.org/10.1145/3290605.3300908>
- Finn, S.M. (2019). Can “Taste” Be Separated from Social Class? *Food Fights*. University of North Carolina Press. doi: 10.5149/northcarolina/9781469652894.003.0005.

- Gargouri, Y., Lefebvre, B. & Meunier, J. (2003). *Ontology Maintenance using Textual Analysis. Unpublished Paper*. Available from: *Montreal: LCIA and LKMDA, UQAM*.
- Grivins, M., Halloran, A. & Kale, M. (2020). Eight megatrends in Nordic-Baltic food systems. *Nordisk Ministerråd*.
- Ihler, A., Newman, D. (2011). Understanding errors in approximate distributed latent Dirichlet allocation. *IEEE Transactions on Knowledge and Data Engineering*, 24(5), pp. 952–960.
- Jin, C., Bouzembrak, Y., Zhou, J., Liang, Q., van den Bulk, L. M., Gavai, A., Liu, N., van den Heuvel, L. J., Hoenderdaal, W., & Marvin, H. J. (2020). Big data in food safety – a review. *Current Opinion in Food Science*, 36:24–32. *Food Safety*.
- Jurafsky, D. (2015). *The Language of Food: A Linguist Reads the Menu*. Reprint edition. New York: W. W. Norton & Company.
- Kāle, M., Šķilters, J. & Rikters, M. (2021). Tracing multisensory food experiences on *Twitter*. *International Journal of Food Design*, 6:2, pp. 181–212.
- Kāle, M. & Agbozo, E. (2021). Utility of large-scale recipe data in food computing. *Baltic Journal of Modern Computing*, 9(2), pp. 155-165.
- Kāle, M. & Rikters, M. (2021). Fragmented and Valuable: Following Sentiment Changes in Food Tweets. Workshop proceedings *Smell, Taste, and Temperature Interfaces (STT21), CHI 2021 Workshop*.
- Kāle, M. & Agbozo, E. (2020a). Tracing complexity in food blogging entries. *CEUR Workshop Proceedings*
- Kāle, M. & Agbozo, E. (2020b). Healthy Food Depiction on Social Media: The Case of Kale on *Twitter*. *DHN Post-Proceedings*
- Khatai, S. *et al.* (2019). An Iterative Model for Text Mining Using Big Data Technology. SSRN Scholarly Paper ID 3356552. Rochester, NY: Social Science Research Network.
- Kim, K.-J. & Chung, C.-H. (2016). Tell Me What You Eat, and I Will Tell You Where You Come From: A Data Science Approach for Global Recipe Data on the Web. *IEEE Access*, 4, pp. 1–1.
- Laguna, L., Fiszman, S., Puerta, P., Chaya, C., & Tárrega, A. (2020). The impact of COVID-19 lockdown on food priorities. Results from a preliminary study using social media and an online survey with Spanish consumers. *Food quality and preference*, 86, 104028. <https://doi.org/10.1016/j.foodqual.2020.104028>
- Loughran, T. & McDonald, B. (2011). When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. *The Journal of Finance*, 66(1), pp. 35–65.
- Lucini, F.R. *et al.* (2017). Text mining approach to predict hospital admissions using early medical records from the emergency department. *International Journal of Medical Informatics*, 100, pp. 1–8.

- Mai, R. & Hoffmann, S. (2015). How to Combat the Unhealthy = Tasty Intuition: The Influencing Role of Health Consciousness. *Journal of Public Policy & Marketing*, 34(1), pp. 63–83.
- Mai, R., Hoffmann, S., Helmert, J.R., et al. (2011). Implicit food associations as obstacles to healthy nutrition: the need for further research. *The British Journal of Diabetes & Vascular Disease*, 11(4), pp. 182–186.
- Marvin, H.J.P. et al. (2017). Big data in food safety: An overview. *Critical Reviews in Food Science and Nutrition*, 57(11), pp. 2286–2295.
- Metcalf, R. (2019). *Food Routes: Growing Bananas in Iceland and Other Tales from the Logistics of Eating*. Cambridge, MA, USA: MIT Press.
- Min, W. et al. (2018). You Are What You Eat: Exploring Rich Recipe Information for Cross Region Food Analysis. *IEEE Transactions on Multimedia*, 20(4), pp. 950–964.
- Min, W., Jiang, S. & Jain, R. (2019b). Food Recommendation: Framework, Existing Solutions and Challenges. *arXiv:1905.06269 [cs]* [Preprint].
- Min, W., Jiang, S. & Liu, L. (2019a). A Survey on Food Computing. *ACM Computing Surveys*, 52(5), p. 36.
- Nag, N., Oh, H., Tang, M., Shi, M. & Jain, R. (2021). Towards integrative multi-modal personal health navigation systems: Framework and application.
- Nag, N., Pandey, V. & Jain, R. (2017). Live Personalized Nutrition Recommendation Engine. *MMHealth'17: proceedings of the 2nd International Workshop on Multimedia for Personal Health and Health Care: October 23, 2017, Mountain View, CA, USA. ACM Workshop on Multimedia for Personal Health and Health Care*.
- Nikolenko, S. I., Koltcov, S. & Koltsova, O. (2017). Topic modelling for qualitative studies. *Journal of Information Science*, 43(1), pp. 88–102.
- Obrist, M. et al. (2019). Space Food Experiences: Designing Passenger’s Eating Experiences for Future Space Travel Scenarios. *Frontiers in Computer Science*, 1, p. 3.
- Obrist, M., Marti, P., Velasco, C., Tu, Y. T., Narumi, T. & Møller, N. L. H. (2018). The future of computing and food: Extended abstract. *Proceedings of the 2018 International Conference on Advanced Visual Interfaces, AVI '18, New York, NY, USA. Association for Computing Machinery*.
- Puerta, P., Laguna, L., Vidal, L., Ares, G., Fiszman, S. & Tárrega, A. (2020). Co-occurrence networks of Twitter content after manual or automatic processing. A case-study on “gluten free”. *Food Quality and Preference*, 86, 103993.
- Rosa, I., Lages, M., Grilo, C., Barros, R. & Guarino, M. P. (2022). mHealth Applications to Monitor Lifestyle Behaviors and Circadian Rhythm in Clinical Settings: Current Perspective and Future Directions. *Frontiers in public health*, 10, 862065.

- Rostami, A., Nagesh, N., Rahmani, A. & Jain, R. (2022). World Food Atlas for Food Navigation. *Proceedings of the 7th International Workshop on Multimedia Assisted Dietary Management*, pp. 39–47.
- Rostami, A., Pandey, V., Nag, N., Wang, V. & Jain, R. (2020). Personal Food Model. *Proceedings of the 28th ACM International Conference on Multimedia (MM '20)*. Association for Computing Machinery, New York, NY, USA, pp. 4416–4424.
- Spence C. (2017). *Gastrophysics: The new science of eating*. Penguin Random House UK.
- Spence, C. (2015a). Visual contributions to taste and flavour perception. *Colour Additives for Foods and Beverages*, Woodhead Publishing, pp. 189–210.
- Spence, C. (2015b). Eating with our ears: Assessing the importance of the sounds of consumption to our perception and enjoyment of multisensory flavour experiences. *Flavour*, 4:3.
- Spence, C. (2021). Explaining seasonal patterns of food consumption. *International Journal of Gastronomy & Food Science*, 24:100332.
- Spence, C. et al. (2016). Eating with our eyes: From visual hunger to digital satiation. *Brain and Cognition*, 110, pp. 53–63.
- Sproģis, U. & Rikters, M. (2020). What can we learn from almost a decade of food tweets. Human Language Technologies – The Baltic Perspective. *Proceedings of the Ninth International Conference Baltic HLT 2020*, IOS Press, 328, pp. 191–198.
- Straka, M. & Straková, J. (2017). Tokenizing, POS Tagging, Lemmatizing and Parsing UD 2.0 with UDPipe. *Proceedings of the CoNLL 2017 Shared Task: Multilingual Parsing from Raw Text to Universal Dependencies. CoNLL 2017*, Vancouver, Canada: Association for Computational Linguistics, pp. 88–99.
- Straka, M., Hajič, J. & Straková, J. (2016). UDPipe: Trainable Pipeline for Processing CoNLL-U Files Performing Tokenization, Morphological Analysis, POS Tagging and Parsing. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16). LREC 2016*, Portorož, Slovenia: European Language Resources Association (ELRA), pp. 4290–4297.
- Svenfelt, A. & Zapico, J. L. (2016). Sustainable food systems with ICT. In *Proceedings of ICT for Sustainability 2016*, pp 194–201. Atlantis Press.
- Thakkar, G. & Pinnis, M. (2020). Pretraining and Fine-Tuning Strategies for Sentiment Analysis of Latvian Tweets. *Human Language Technologies—The Baltic Perspective—Proceedings of the Ninth International Conference Baltic HLT 2020*; IOS Press: Kaunas, Lithuania, 2020; pp. 55–61
- The Lancet Gastroenterology & Hepatology. Obesity: another ongoing pandemic. *Lancet Gastroenterol Hepatol* 2021; 6: 411.
- Tran, B. X., Nghiem, S., Sahin, O., Vu, T. M., Ha, G.H., Vu, G. T. & Ho, C. S. (2019). Modeling research topics for artificial intelligence applications in medicine: latent Dirichlet allocation application study. *Journal of medical Internet research*, 21(11), e15511.s

- Twitter Developer Platform (n.d.a). Consuming streaming data. <https://developer.twitter.com/en/docs/tutorials/consuming-streaming-data>. Accessed 15 August 2021.
- Vasiloglou, M. F., Lu, Y., Stathopoulou, T., Papathanail, I., Faeh, D., Ghosh, A., Baumann, M. & Mougiakakou, S. (2020). Assessing mediterranean diet adherence with the smartphone: The medipiatto project. *Nutrients*, 12(12).
- Velasco, C., Michel, C. & Spence, C. (2021). Gastrophysics: Current approaches and future directions. *International Journal of Food Design* 2021;6:137-52.
- Velasco, C. & Obrist, M. (2020). *Multisensory Experiences*, Oxford University Press.
- Vīksna, R. & Skadiņa, I. (2020). Large Language Models for Latvian Named Entity Recognition. In Utka, A. et al. (Eds) *Frontiers in Artificial Intelligence and Applications*. IOS.
- Wang, Q. J. et al. (2019). The Role of Intrinsic and Extrinsic Sensory Factors in Sweetness Perception of Food and Beverages: A Review. *Foods (Basel, Switzerland)*, 8(6), p. E211.
- Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., Garnett, T., Tilman, D., DeClerck, F., Wood, A., Jonell, M., Clark, M., Gordon, L. J., Fanzo, J., Hawkes, C., Zurayk, R., Rivera, J. A., De Vries, W., Majele Sibanda, L., Afshin, A., Murray, C. (2019). Food in the Anthropocene: the EAT-Lancet Commission on healthy diets from sustainable food systems. *Lancet (London, England)*, 393(10170), pp. 447–492.
- Xu, X., Sharma, P., Shu, S. et al. (2021). Global greenhouse gas emissions from animal-based foods are twice those of plant-based foods. *Nat Food* 2, pp. 724–732.
- Yamakata, Y., Mougiakakou, S., & Jain, R. (2022). Cea++2022 panel - toward building a global food network. In Proceedings of the 1st International Workshop on Multimedia for Cooking, Eating, and Related Applications, CEA++ '22, pp. 59–60, New York, NY, USA. Association for Computing Machinery.

Pateicības

Šajā piecu gadu ilgajā doktorantūras studiju procesā vēlos pateikties diezgan daudziem cilvēkiem, kas uzmundrinājuši, iedrošinājuši un nav ļāvuši ilgstoši skumt par studiju un pētniecības procesa, kā ierasts, grumbulaino un līkumoto ceļu.

Vislielākā pateicība pienākas maniem vecākiem Askoldam Kālim un Vijai Kālei, kuriem pateicoties pasaules izziņas process man šķiet interesants, aizraujošs un skaists.

Promocijas darba izstrādes procesā liela nozīme bijusi maniem līdzautoriem Ebenezeram Agbozo un Matīsam Rikteram, kas nu jau kļuvuši par draugiem un domubiedriem – un draudzība ir pats labākais pētniecības procesa blakusefekts. Vislielākā pateicība profesoram Ramešam Džeinam (Ramesh Jain) par pētniecības ceļa izgaismošanu un ieviešanu brīnišķīgā starptautisku pētnieku kopienā, kas ļāvis just to neatkārtojamo būšanu epicentrā sajūtu un pietuvošanos jauniem, patiešām nozīmīgiem atklājumiem. Liels paldies manam promocijas darba vadītājam profesoram Jurģim Šķilteram un manai darba vietai – Ziemeļvalstu Ministru padomes birojam Latvijā par doto iespēju studēt doktorantūrā.

Vēlos pateikties arī citiem LU Datorikas fakultātes mācībspēkiem – profesoriem Kārlim Podniekam, Jurim Borzovam un Andrim Ambainim par izcilajiem kursiem un semināriem studiju procesā. Jo īpaši vēlos pateikties profesoram Leo Seļāvo – paldies par nenogurstošu atbalstu, iedvesmu un augstajiem standartiem darbā un dzīvē. Nesaucot vārdos – vislielākais paldies maniem draugiem par iespēju kopā par visu pasmaidīt.

Paldies manam brālim Mārtiņam Kālim, kas palīdzēja un motivēja uzsākt studijas doktorantūrā līdz ar viņu. Lielu pateicību esmu parādā arī savam vīram Younes Filali, kurš uz saviem pleciem iznes ikdienas rūpes par ēst gatavošanu un pārtikas menedžmentu mūsu ģimenē, līdz ar to laiku šo rūpju vietā varēju veltīt studijām un pētniecībai. Paldies meitām Selmai un Munai, kuras vienmēr uzmundrināja un pat piedalījās dažādos mācību procesos, piemēram, atvērtā koda programmatūras “Blender” apguvē studiju kursā pie profesora Leo Seļāvo, kas vēlāk izauga par “Blender” 3D modelēšanas entuziastu klubiņu maniem draugiem un viņu bērniem.

Studiju procesu neaizmirstamu padarīja arī LU Datorikas fakultātes metodiķes Anitas Ermušas (lai viņai vieglas smiltis) sirsnība, tāpat arī jaukā komunikācija ar Promocijas padomes sekretāri Rutu Ikaunieci. Visbeidzot, vislielākais paldies manis vadīto kursu studentiem – jūsu atsauksmes es pārlasu ikreiz, kad vajag uzmundrinājumu dzīvei. Paldies!