МЕЃУНАРОДЕН СЛАВЈАНСКИ УНИВЕРЗИТЕТ

ФАКУЛТЕТ ЗА **ТЕХНИЧКИ НАУКИ И ИНФОРМАТИКА**
НАСОКА: **КОМПЈУТЕРСКИ СИСТЕМИ И МРЕЖИ**

МАГИСТЕРСКА РАБОТА

ИНТЕЛИГЕНТНИ СИСТЕМИ СО ИНТЕГРИРАНА ГОВОРНА КОМПОНЕНТА НА ВЕШТАЧКА ИНТЕЛИГЕНЦИЈА

Ментор:
Проф. Д-р. Сашо Манасов

Изработил:
Милан Велкоски

Maj, 2024
Св. Николе

ВОВЕД

Светот денес стана исклучително предизвикувачко место за живеење со сé повеќе и повеќе штетни фактори кои ја загрозуваат нашата благосостојба. Додека модерната медицина брзо напредува, предизвиците во здравството се појавуваат со стареење на населението, менталното здравје и испорака на здравствена заштита. Свеста за овие големи прашања од нашата ера почнува да го привлекува јавното мислење, а таканаречените “паметни технологии“ брзо се шират за да одговорат на предизвиците на општеството на 21 век.

Според Transparency Market Research, пазарот на дигитално здравје ќе се зголеми на половина милијарда долари до крајот на 2025 година. Во ова сценарио, луѓето ќе имаат зголемена побарувачка за медицински помагала за превенција, следење и асистирана дијагноза на болести и клинички нарушувања, како и за грижа или континуирана самопроверка за одржување на здрава состојба. Дигиталната здравствена заштита, достапна од паметни уреди, ќе стане главен тренд, особено во хаотичен свет каде што повеќе луѓе сакаат лесен и моментален одговор без потреба од посета на лекар за рутински прегледи. Посебен интерес ќе биде за оние технологии кои би можеле да ја заклучат нашата состојба директно од нашиот паметен телефон. Гласот е многу директен и непосреден сигнал за користење, и е богат со информации за говорникот: говорна содржина, емоции, свесност и важни биомаркери.

1. Емоционална вештачка интелигенција

Мултисензорска анализа на емоции

Создавање мерливи емоции: Интелигентна анализа на мултимодални податоци

Несвесни емоционални реакции (Unconscious emotional reactions) кои се предизвикани од надворешни впечатоци, имаат големо влијание врз тоа:

- како ја перципираме нашата околина, и
- како донесуваме одлуки.

Овие процеси се сложени и се случуваат потсвесно и затоа тешко се откриваат. За да се опишат и разберат овие свесни и несвесни реакции, емоциите во психологијата се опишани во три категории:

- **Физички реакции**: на пример, забрзано чукање на срце или тенок слој пот на кожа.
- **Субјективно искуство**: како личноста ја перцепира ситуацијата врз основа на индивидуално искуство.
- **Однесување**: на пример, насмевка или нагласено (стресно) пишување на тастатура.

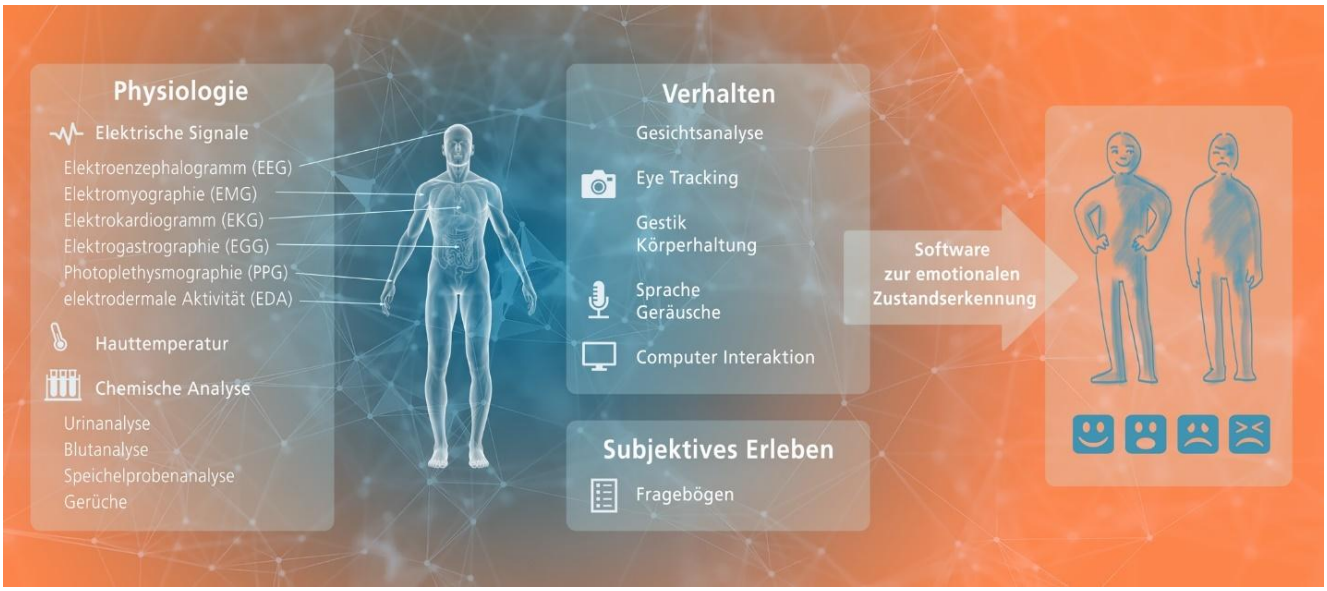
Цел на истражувачите е **да ги анализираат** емоционалните реакции што е можно похолистички и **објективно да ги толкуваат** преку интелигентно оценување на овие три перспективи.

Врз основа на ова, може да се развијат **човеч-технологија интерфејси** кои:

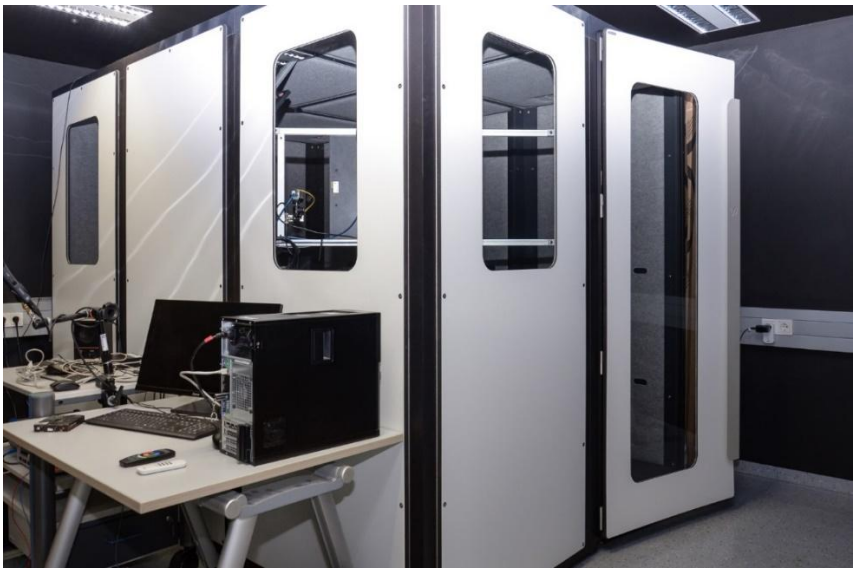
- можат да ја препознаат емоционалната состојба на корисникот, и
- соодветно да реагираат.

За таа цел, се спроведуваат истражувања за мултимодална анализа на психо-физиолошка реакција на корисникот – за подобро разбирање на несвесните реакции на корисниците преку сеопфатна анализа и визуелизација на нивните емоционални состојби:

- Прилагоден дизајн на студија за корисници.
- Висока точност преку прецизно снимање на податоци.
- Мултисензорска фузија на податоци.
- Анализа на мултимодални множества на податоци базирана на вештачка интелигенција.
- Транспарентност преку пренос на целосен сет на податоци вклучувајќи анализа (без “црна кутија“).



- **Снимање податоци**: Висококвалитетни податоци се основа за класификација на емоции користејќи алгоритми базирани на вештачка интелигенција. За да се направи ова, се развиваат научно издржани дизајни на студии кои се приспособени токму на поединечно прашање. Врз основа на ова, се избира група на испитаници што ја одразуваат целната група и се дефинираат мултимодални модалитети за снимање на податоци.
- **Брзо и флексибилно поставување студија**: мултимодална мерна кабина овозможува ефикасно поставување на истражување со синхронизирано снимање на податоци. Во зависност од барања и желби на клиенти, може да се додадат специфични мерни модалитети.
- **Сигурно (Reliable) препознавање на емоции**: мултимодалните податоци се снимаат синхроно, така што **стимулот** и **реакцијата** можат прецизно да се мапираат еден на друг. Преку спојувањето на различните влезни податоци засновано на вештачка интелигенција, се добива посеопфатна и попрецизна слика за тест состојба на испитаниците. Можност за 360° анализа на емоции во затворена мерна кабина без пречки/шумови (interference-free).



- **Брза евалуација на податоци**: Линијата на податоци овозможува анализа на претходно обработени податоци за да може да се постигне ефикасна постобработка на податоци и брзи резултати.
- **Високо ниво на безбедност на податоците**: Се посветува големо значење на заштита на личните права на испитаниците, како при снимање на податоци, така и за време на анализа. За да се направи ова, се земаат во предвид сите неопходни етички барања и прописи за заштита на податоци.
- Се користи софтвер за анализа на емоции базирано на камера, со кој се пренесуваат само анонимизирани метаподатоци.

Симулатор за возење: Анализа на емоции прилагодена на автомобилската средина. Едноставно и брзо снимање на податоци за задачи за возење и сообраќајни ситуации. Симулатори за возење, целосно опремени со камери, осветлување и системи за мултимодално снимање биосигнали. Медицински референтни системи и мрежа на медицински експерти. Индивидуална адаптација на тест поставката кон барањата во автомобилската средина. Оптимално и без проблеми модално поставување. Инфраструктура за усогласеност со барања за доверливост.



Мултимодална фузија на податоци

Голем број достапни сензори за психо-физиолошки мерења нудат широк опсег на можности за снимање на емоционални состојби. Предизвик лежи во избор, мерење и спојување на вистински сигнали. За да се направи ова, се комбинира знаење за **длабинско учење** (deep learning) и **невронски мрежи** со методи на филтрирање за спојување на информации. На овој начин, може да се искористи целосниот потенцијал на мултимодалноста за споредување на индивидуални проценки. Целта е да се оптимизираат модели за предвидување преку фузија на податоци за да може да се постигне оптимална класификација на афективни состојби.

Друга предност на мултимодална анализа на податоци е нејзината робусност во однос на неуспешни поединечните влезни параметри, особено во безбедносно критични апликации. На пример, пулс може да се одреди со помош на паметен часовник и во исто време со помош на решение базирано на камера. Алгоритмот му верува на оној модалитетот кој има најмала мерна несигурност. Ако еден од двата параметри не успее, сепак се добиваат сигурни (reliable) резултати, дури и ако квалитетот на предвидувањето повеќе не е оптимален.

2. Анализа на израз на лице за медицински апликации

Како е пациентот?

Што влијае на него и колку?

Ова се централни прашања на секоја медицинска историја. Изјавите на пациентот, како и неговиот хабитус се користат за да се процени сериозноста на неговото заболување. Но, што ако пациентот нема способност вербално да артикулира, било поради возраста или медицинска позадина (background)?

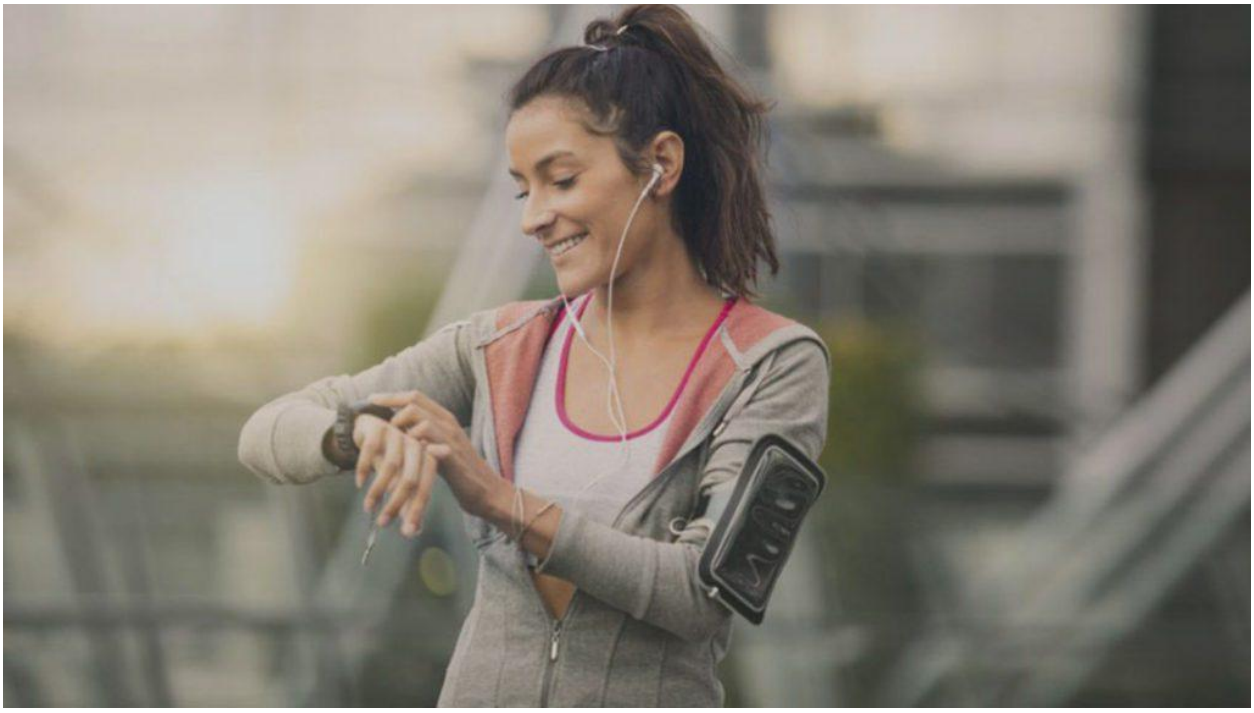
Дигитализацијата која брзо напредува нуди нови можности во медицинската технологија. Превенцијата, дијагнозата и терапијата може суштински да се подобрат преку користење на вештачка интелигенција и нови стратегии за интеракција меѓу човекот и технологијата. Лекари на иднината, исто така, ќе можат брзо и врз основа на податоци да разменуваат идеи низ дисциплини. Тие се поддржани со однапред подготвени податоци. Пациентите имаат корист од брзи дијагнози и концепти за терапија приспособени конкретно за нив.

SHORE® софтвер за откривање и анализирање на изрази на лице, може да поддржува процедури на лице место и процедури за дигитална анамнеза. Особено во медицината, системските одлуки мора да бидат транспарентни, цврсти и разбирливи за да може да се даде најголема можна доверба во резултати генерирани на овој начин и да се обезбеди дијагностичка поддршка на клиничките специјалисти.

Денешните истражувачки полиња моментално се фокусираат првенствено на откривање на емоции кај пациенти од спектар на аутизам, робусно откривање на акциони единици како основа за откривање болка, вмрежување на различни податоци базирани на сензори и следливост на одлуки во невронски мрежи.

Како технологиите за анализа на говор (глас) ќе го променат нашето искуство за благосостојба и здравствена заштита

Благосостојба: предизвик во светот денес



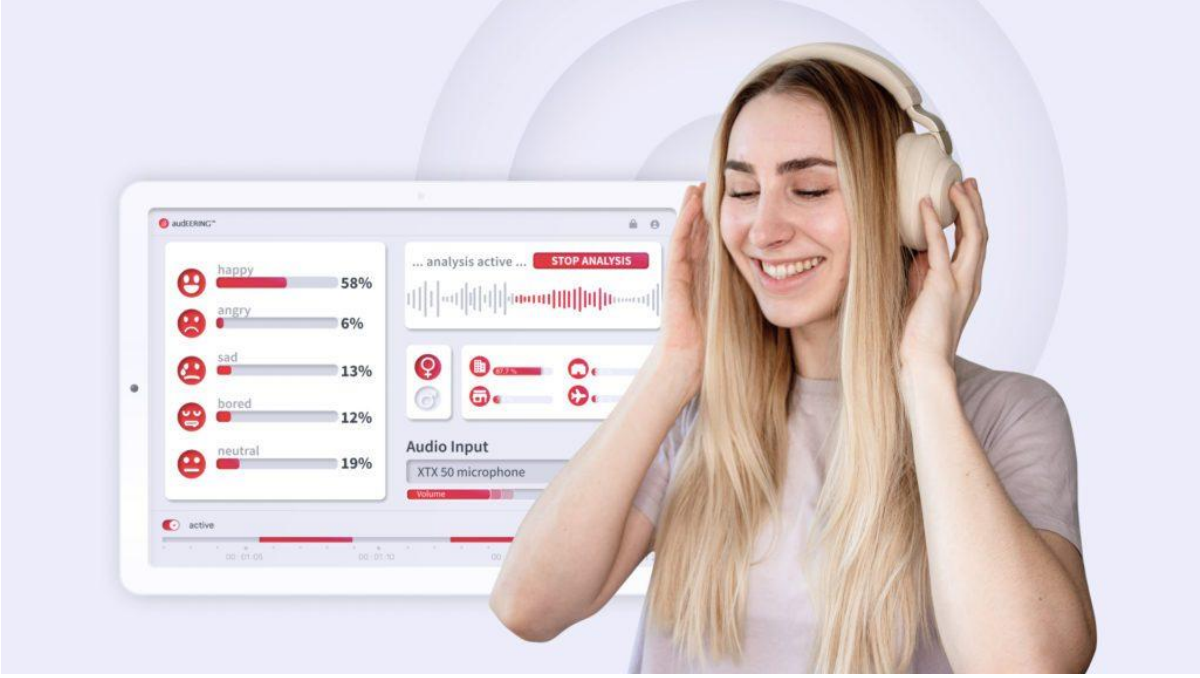
Што е посебно во гласот надвор од говорот?

Некои од најтешките невролошки заболувања се дегенеративни: раното откривање и проценка на прогресијата на болеста се исклучително важни фактори за да се олесни ефективно лекување на овие болести. audEERING, како водечка компанија во интелигентна аудио обработка и паралингвистичка анализа, обезбедува технологија способна да идентификува пациенти со рани симптоми на Паркинсонова болест преку анализа на суптилни промени во гласот и постигната 92% точност на класификација за здрави и заболени примероци. Во овој контекст, гласот може да обезбеди и некои насоки за пациентите и да им помогне да ги приспособат лековите поефикасно и на индивидуализиран начин. Гласот може да се користи како моќен индикатор не само за Паркинсоновата болест, туку и за мултиплекс склероза, Хантингтонова и Алцхајмерова болест за да се овозможи подобра дијагноза и препораки за третман. Покрај тоа, постојат објавени докази дека отчукувањата на срцето може да се заклучат од говорните искази, и затоа гласот содржи уште еден важен биомаркер кој потенцијално би овозможил далечинско оценување и следење на пациентот. Освен ова, гласовните технологии можат да помогнат и во напорите за рехабилитација. Во неодамнешните студии,

јазичните оштетувања по мозочен удар беа успешно третирани со помош на гласовна терапија за говор и јазик. Дополнително, гласот може да се користи за да се зачува нашата благосостојба и да се препорачаат подобри одлуки како навики за здрава исхрана. Всушност, почнувајќи од деведесеттите, бројот на деца со прекумерна тежина во Германија е двојно зголемен, при што приближно еден од шест адолесценти има BMI над 25. Во таа перспектива, auDEERING исто така е активен во истражување на модели на звуци за да предвиди можна медицинска состојба како дебелина или булимија.

Иновативни паметни технологии за здравствена заштита

Во здравството и благосостојбата, постојат апликации за предвидување симптоми или помош за дијагностицирање на различни болести, од когнитивни нарушувања до срцев удар. Неколку стартапи користат гласовни технологии за да помогнат во подобрувањето на животот на оние со тешкотии во говорот и/или слухот, со потенцијални иновации за следење на прогресијата на болеста со текот на времето. Другите апликации се фокусираат на помошта за стари лица, помагајќи им на постарите да се поврзат со семејството, да добијат пристап до услуги, да усвојат здрави навики и да напредуваат самостојно. Ова е очигледно ако размислиме за тоа како паметните телефони во денешно време се буквално во секој џеб, што го отвора патот за широка пристапност до гласовни технологии. Неодамна, постигнати се огромни технолошки достигнувања, а особено за вештачката интелигенција, работите станаа интересни. Huawei HONOR 10 има посебна единица за невронско процесирање, на која операциите за процесите на машинско учење можат ефикасно да се извршуваат. Pixel 2 на Google и iPhone X имаат слични технологии. Ова може да ги поттикне технологиите за здравствена заштита во голема мера, бидејќи анализа на чувствителни податоци од корисниците може да се извршува локално на телефон. Особено во доменот на здравствената заштита, безбедноста на податоците е чувствителна тема, а локално обработени податоци се помалку подложни на напади од податоци со кои се ракува од облак услуги.



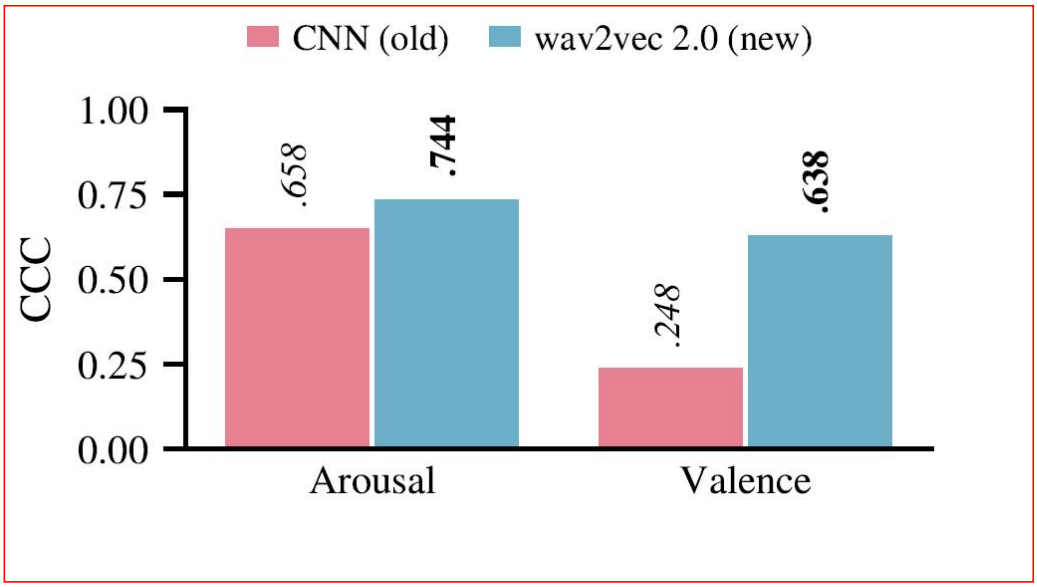
Затворање на валентниот јаз во препознавање емоции

2021 година е возбудлива година за иновативни истражувања кои работат на препознавање на емоции од гласот. Користејќи ги придобивките од неодамнешниот напредок во компјутерска архитектура, за прв пат се изградија модели кои предвидуваат **валентност** (т.е. колку негативно или позитивно звучи еден исказ) со слична висока прецизност како **возбудување** (т.е. колку возбудено звучи еден исказ). Затворањето на валентниот јаз е голем предизвик во **препознавање на говорни емоции** (SER-speech emotion recognition). За разлика од претходни пристапи, овие модели:

- ги изведуваат своите предвидувања директно од необработено аудио, и
- не бараат транскрипции на текст.

Валентно предвидување без користење на лингвистички информации

Реткоста на означени податоци (The sparseness of labeled data) е една од најголемите бариери во истражување на SER. За да се надмине ова ограничување, денес фокусот се префрла од специјализирани архитектури обучени за дадена задача кон модели врз основа за општа намена.



Слика. Корелација помеѓу основната вистина и предвидувањата добиени со специјализирана архитектура на CNN (стара) наспроти моделот wav2vec 2.0 базиран на основа (нов) на MSP-Podcast (повисоко е подобро).

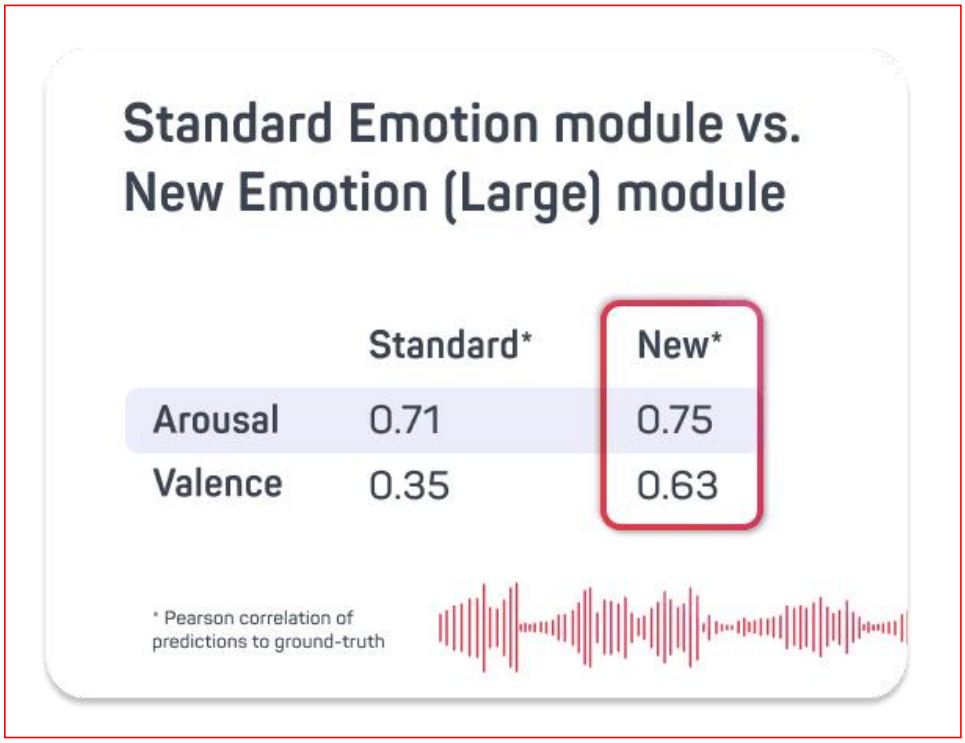
Ваквиот иновативен систем добива врвни перформанси за **валентно предвидување** без употреба на експлицитни лингвистички информации, со **коефициент на корелација на усогласеност** (**CCC**-concordance correlation coefficient) од 0,638 на MSP.

devAIce TM – аудио анализа за хардвер и софтвер

devAIceTM е основната технологија на audEERING на која расте секоја понатамошна идеја. Дизајниран е да **препознава емоции**, акустични сцени и многу други функции од аудио. И во реално време и за кој било производ. Моделите со вештачка интелигенција работат солидно дури и со ограничена моќност на процесорот. Основната технологија на audEERING може да следи над 6.000 гласовни параметри во реално време. Истражувачи користат врвни методи на вештачка интелигенција како длабоки невронски мрежи и учење без надзор (unsupervised learning).



Споредувајќи ги стандардните модели на емоции со новите модели се гледа драматично зголемување на точноста на предвидувањето на валентноста и забележително зголемување на проценката на возбудата:



Спој на емоции од говор и текст

Самостојни модели за препознавање емоции од аудио и анализа на чувства од англиски текст веќе се достапни во devAIceTM. Со новото издание на devAIceTM, се воведува нов Multi-Modal Emotion модул. Тој нуди посветен, единствен модел кој прифаќа аудио и текст како влез и ги спојува податоците во едно комбинирано предвидување на емоции.

Анализи насочени кон корисниците, каде што говорот на само еден специфичен говорник треба да се обработи, а сите други гласови да се игнорираат.

Интонација

Се користи за да се процени колку едно лице зборува монотоно или живо.

Целосниот потенцијал на технологија за анализа на говор со вештачка интелигенција го користи гласот како природен ресурс за создавање вредност за луѓето и нивниот секојдневен живот.

Тековни трендови за апликации базирани на говор (глас)

Дали некогаш сме се запрашале како би било да имаме личен асистент да нé следи каде и да одиме? Секој што поседува паметен телефон или персонален компјутер мора да стапил во контакт со некој од многу гласовни асистенти: Siri, Cortana, Google (најпознати). Нивниот напредок во разбирањето на она што го кажуваме е запрепастувачки во последните неколку години. Тие беа во можност да го постигнат тоа поради напредокот што полето наречено **Автоматско препознавање на говор** (ASR- Automatic Speech Recognition) го забележа само во оваа деценија.

Меѓутоа, иако перформансите на ASR постојано се подобруваат, а гласовните асистенти постојано се подобруваат во разбирањето на она што го кажуваме, често се случува да не разберат што мислиме. Поради тоа треба да се повтори нешто повеќе пати, или да се поедностави изразот, да се ограничи употребата на зборови, да се скратат реченициите итн. Следна бариера за целосно разновидни интелигентни разговори со машина е **обработка на природниот јазик** (NLP- Natural Language Processing), накратко автоматската анализа на нашата реченица.

2018 година е наречена година на NLP од многу водечки истражувачи на вештачка интелигенција - и тоа со добра причина. Тоа е година што забележа огромен напредок во користењето на длабоки невронски мрежи за решавање на многу од задачите што ја предизвикуваат истражувачката заедница со децении. Ова ја отвора вратата за поинтелигентни апликации за гласовен асистент. Но, она што лежи надвор од тоа може засекогаш да го промени начинот на кој го замислуваме светот околу нас.

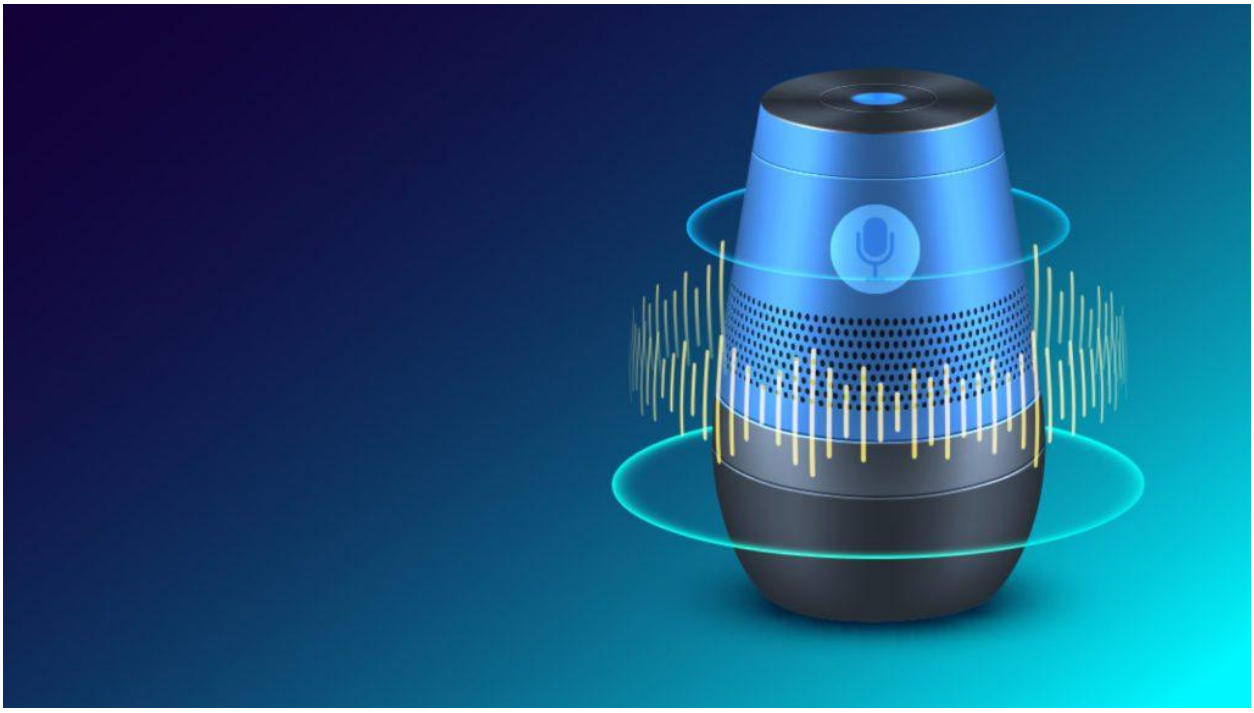
Како комуницираме денес со повеќе наши електронски уреди?

За да добиеме одговор од нашиот компјутер или паметни телефони, мора да пишуваме со нашата тастатура, да кликнеме со глувчето, да скролираме, да се движиме со табовите, да поминеме со прстот, да туркаме со прстот и така натаму. Сега да размислиме како комуницираме со други човечки суштества. Големи се шансите, поголем дел од времето едноставно да разговараме со нив.

Говорниот јазик е нашиот избран тип на комуникација уште од нашите рани денови како вид. Всушност, веројатно е многу подлабоко во нашата еволутивна историја од првите луѓе кои се појавија. Звуците се толку длабоко вкоренети во начинот на кој комуницираме со нашиот свет што е невозможно да се размислува за иднина во која нашата комуникација со машини не е барем делумно заснована на овој модалитет.

И моментот за тоа конечно дојде. ASR беше првата граница. Очигледно, ако машината не може да го одбере нашиот избор на зборови, тогаш нема смисла да се обидуваме да комуницираме со неа преку говор. Сега, сепак, нашите телефони се способни да препишуваат речиси сè што ќе кажеме веднаш, па дури и под сосема различни услови на бучава.

NLP беше втора бариера. Каква корист беше нашата машина да запише сè што кажуваме, ако речиси ништо немаше никаква смисла за тоа. Се разбира, ние сме далеку од алгоритам кој може да го разбере човечкиот јазик со сите негови сложености и нијанси, но тоа не е она што ни треба. Најважно е основната употребливост.



Има многу повеќе пресвртници за вистинска разговорна вештачка интелигенција, но двете најважни се скршени. Ќе се случи гласовно контролиран оперативен систем (voice controlled operating system - се случува. Siri, Alexa, Google, Cortana). Сите тие го менуваат начинот на кој комуницираме со нашите уреди во секојдневниот живот. Наскоро тие ќе станат норма. Ќе биде незамисливо да поседуваме машина која не може да разбере што зборуваме, слично како што е речиси незамисливо да поседуваме машина што не може да ја контролираме со допир на прст. Слично како што е незамисливо да поседуваме машина која е достапна само преку тастатура - и терминал.

Човек во јамка (циклус) – Како се создава вештачка интелигенција

Вештачката интелигенција се потпира на луѓето. Развивајќи технологијата за вештачка интелигенција како што е audEERING, треба да се разбере човечката перцепција. Затоа се работи со обучен внатрешен тим за прибелешки. Тоа се луѓе различни по возраст, пол, потекло и сегашност, кои обезбедуваат знаење за нова технологија за вештачка интелигенција. Секојдневната перцепција ни овозможува

да ја сфатиме емоционалната состојба на нашиот комуникациски партнер во различни ситуации. Во процесот на човечко машинско учење треба да се даде суштински влез на алгоритмот. Фокус треба да се даде на процесот на учење на човек-машина.

Ставање на луѓето во средина

Секојдневната перцепција ни овозможува да ја сфатиме емоционалната состојба на нашиот комуникациски партнер во различни ситуации. Најчесто тоа се случува несвесно. Еден типичен пример е повик со добар пријател. Го прашуваме нашиот пријател: “Како си?“ и одговорот може да биде: “Добро сум“. Сега да замислиме дека нашата реакција е: “О, што не е во ред? Дали сакаш да разговараш?“ Како звучи личноста додека користиме позитивна содржина, што нé тера да реагираме на овој начин?

Создавање несвесно свесно

Во процесот на човечко машинско учење (Human Machine Learning) треба да му дадеме суштински влез на алгоритмот. Основа е огромна количина на податоци за обука. Податоците што се користат за VoiceAI се аудио. Имајќи ги овие податоци достапни, следниот чекор е да знаеме што сакаме да дознаеме од овие податоци. Колку е појасна целта, толку појасно може да биде прашањето. Дадените одговори се влез што се користи за обука на вештачката интелигенција. Во моментот на одговор, прибележувачот се обидува да сфати на што се фокусира неговата или нејзината перцепција. Ова е моментот кога несвесното станува свесно.

Ефект на Карпентер

Врз основа на ефектот на Карпентер, слушањето груб глас може да предизвика гребење во грлото. Карпентер ефектот го објаснува внатрешното разбирање на физиолошките движења на друго лице со нивно имитирање несвесно. Во однос на областа на логопедијата, овој ефект се користи за да се доловат тенденции на болести на гласот на една личност. Понатаму, гласот ги пренесува луѓето како емпатични суштества кои се достапни за него. Затоа можеме да

направиме разлика помеѓу содржината на нашиот пријател и начинот на зборување.

Вештачка интелигенција како огледало на човечката перцепција

Како може да се случи вештачката интелигенција да стане дискриминаторска, на пример во однос на полот и потеклото? Технологијата за вештачка интелигенција учи од дадени извори на податоци и нејзини ознаки. Ние како луѓе правиме разлики во нашиот секојдневен живот помеѓу различни јазици:

- поради посебни јазични појави,
- во однос на структурата на формулацијата на јазикот, и
- фонетскиот звук.

Се разликуваме и помеѓу женски и машки начин на зборување. Луѓето се социјализирани суштества, на некој начин обучени да интерпретираат посебни атрибути. Дополнително, постојат реални физиолошки разлики помеѓу машкиот и женскиот вокален тракт што влијае на акустичното создавање на гласот.

Улогата на auDEERING како водечки светски иноватор во Voice AI

Развивањето на вештачката интелигенција со одговорност кон човештвото, науката и транспарентноста во технологијата значи да се разберат податоците. Во суштина, тоа значи **разбирање на човечката перцепција**.

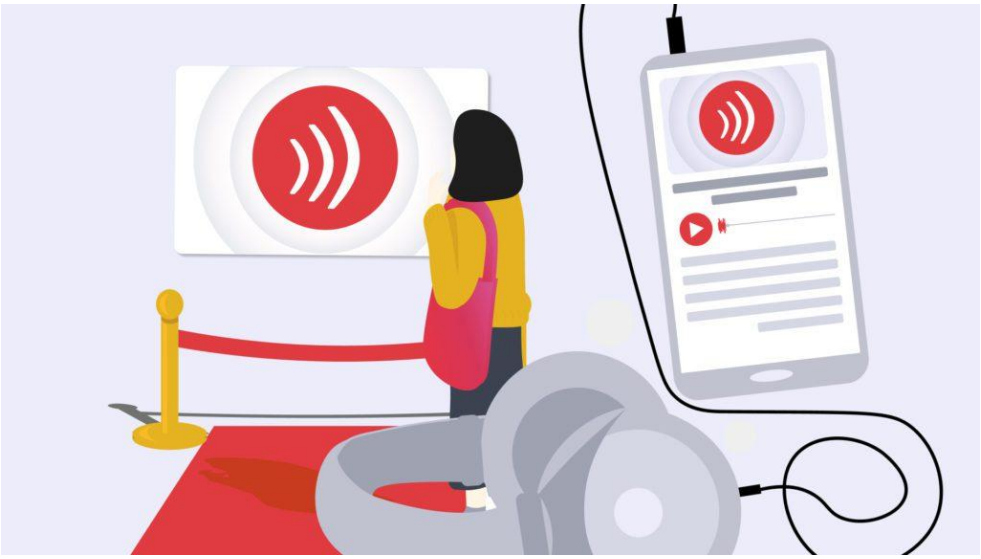
- Како луѓето го сфаќаат она што го слушаат?
- Кои лични информации во позадина влијаат на нашата човечка перцепција?
- За да се одржи различноста, мора да се создаде различност.

Сепак, ова не може да се направи еднодимензионално, туку мора да се долови на сите нивои. Соодветно на тоа, развојот на вештачката интелигенција може да значи додавање вредност што не ја има само технологијата. Тоа значи:

- **да се добие увид во сложеноста на човечкиот ум и перцепција**.

Но затоа комплексноста на човечкиот ум е незаменлива.

Проект (SHIFT) за трансформирање на културното наследство



SHIFT: Метаморфоза на културно наследство во зголемени хипермедијални можности. Стратегија за дигитална трансформација и прифаќање алатки во креативни и културни индустрии (CCI - creative and cultural industries), каде што напредокот заостанува.

audEERING учествува во проектот SHIFT со цел да **синтетизира емоционални говорни описи на историски експонати**. Ова ќе го промени начинот на кој се доживува историски споменик, особено за лицата со оштетен вид.

Конзорциум на водечки светски организации

Конзорциумот SHIFT обединува 13 водечки светски истражувачки, индустриски и мали и средни претпријатија, координирани од SIMAVI (Романија), со визија за зајакнување на влијанието на средствата од културното наследство. Ќе ја зголеми привлечноста на историските артефакти, ќе ја подобри нивната пристапност и употребливост за секого преку подобра застапеност на содржината, збогатени кориснички искуства, инклузивен дизајн, обука и поангажирани бизнис модели.

Проектот е стратешки замислен да испорача, преку унапредување надвор од најсовремената технологија, збир на лабаво поврзани технолошки алатки кои им нудат на институциите за културно наследство поттик да го стимулираат растот и да им пренесат нови искуства на сите граѓани (вклучувајќи и луѓе со попреченост) со прифаќање на најновите технолошки иновации во:

- вештачка интелигенција,
- машинско учење,
- мултимодална обработка на податоци,
- методологии за трансформација на дигитална содржина,
- семантичка претставување,
- лингвистичка анализа на историските записи
- употреба на haptics интерфејси.

Меѓу други клучни резултати ќе бидат:

- алатки и алгоритми за ревитализација на историски и културни содржини со висока вредност;
- засилено зачувување на историските архиви со користење на современи јазични модели;
- придонес кон меѓународни стандарди за размена на модели на метаподатоци со културни институции, и
- заштита на авторски права на сопственоста на содржината.

Кому е насочена SHIFT?

Резултатите од SHIFT ќе бидат насочени кон B2B кориснички групи, вклучувајќи институции на културно наследство (архиви, библиотеки, музеи), креатори на дигитални содржини, забава и креативна и културна индустрија, како и крајни корисници како што се посетители на музи, веб-заедници и ранливи групи.

Планирањето на проектот AGILE ќе усвои цикличен пристап за вклучување со заедници на засегнати страни и ќе воспостави методологија за континуирана евалуација, истовремено развивајќи неколку случаи на употреба и демонстрации. Стратегии за дисеминација и комуникација на SHIFT ќе промовираат широко

усвојување на неговите резултати низ културните и креативните индустрии. Плановите за идна експлоатација ќе се фокусираат на користење на современи пристапи на бизнис модел за SHIFT и неговите компоненти.

Емотивниот влез на audEERING во SHIFT

audEERING's со своето долгогодишно искуство во афективната гласовна анализа управувана од вештачка интелигенција учествува во проектот SHIFT со цел да се синтетизираат емоционални говорни описи на историски експонати. Ова ќе го промени начинот на кој се доживува историски споменик, особено за лицата со оштетен вид.

Целта за извонредна виртуелна обиколка на историски локалитети е акустична реконструкција на настани што се случиле во тој споменик на култура. Технологијата за емоционална синтеза на говор на audEERING е клучен елемент за **виртуелна реконструкција на историските битки** и **вистински моменти од животот во минатото**.

Емоционализација на синтеза на говор станува сé повеќе важна вредност на работата на audEERING, благодарение на знаењето засновано на истражување што се создава за емоционално изразување на гласовите.

Препознавање емоции за терапија со Аутизам

Децата и возрасните со аутизам понекогаш имаат потешкотии да ја изразат и интерпретираат сопствената и туѓата емоционална возбуда. Во најлош случај, тоа може да доведе до социјална изолација, во поблаги случаи до многу непријатни ситуации засновани на недоразбирања.

Со години, психолозите работат со аутистични деца и интерактивна технологија и развија игри кои им помагаат на овие деца да ги обучат своите способности за разбирање емоции.

Заеднички проект на повеќе универзитети и компании има за цел да развие платформа за роботи која може да се користи во тераписки сесии за да овозможи децата разиграно да го истражуваат и трениратсветот на емоционално изразување.

За да им помогне на децата да го проценат ефектот од нивниот вокален израз, audEERING сними база на податоци за недијагностицирани деца што може да послужи како референца за општествено прифатливи изрази на возбуда и валентност. Оваа база на податоци, платформа на audEERING, се користи за обука на класификатори на длабоки нервни мрежи за откривање на возбуда и валентност во говорот. Вака, децата можат да добијат повратна информација од роботот. Роботот не мора да се потпира само на анализа на говор бидејќи аудио предвидувањата ќе се толкуваат заедно со излезот од модулите за анализа на изрази на лице (мимики) и срцеви отчукувања.

Тие првично ќе се комбинираат во пристапот на **доцна фузија**. Доцна фузија значи дека резултатите од секој модалитет (мимика, говор и срцев ритам) ќе се комбинираат на крајот за да се постигне еден конечен резултат (комбинација може да се направи со пристап на машинско учење). Ова има предност што за секој модалитет класификаторите може да се обучуваат на сопствена база на податоци. Исто така се истражуваат пристапи за рана фузија каде што влезните сигнали или извлечените параметри од секој модалитет ќе се комбинираат во еден единствен влез за класификатор штом се достапни заеднички бази на податоци.

Понатаму, audEERING развива софтверска компонента која го менува говорниот излез на роботот за да звучи поемотивно привлечно и да ја подобри природноста со додавање варијации на **прозодија**, т.е. **мелодија** и **ритам** на говорот. Интеракцијата со роботот на тој начин е поинтересна и ќе ја зголеми веродостојноста на роботот како интелигентен ентитет компетентен за емоции, и интеракцијата се потпира на алгоритми за вештачка интелигенција.

Препознавање и перцепција на емоционален излез е суштински дел од човечката комуникација. За да се развијат **социо-емоционални комуникациски вештини** на деца со аутизам, терапијата треба да се фокусира на тоа. Во проектот ERIK се развива и тестира нова форма на терапија со помош на роботски систем.

Користење на роботи во терапија со деца

Користење на роботи се покажа како корисно во терапијата на овие деца. Прво, затоа што овие деца роботот го доживуваат како помалку сложен. Затоа е попредвидливо и помалку застрашувачко. Второ, затоа што децата со аутизам често покажуваат голем интерес и многу силни страни на техничка област. Роботот ги детектира изразите на лицето и говорот во интеракција со детето. Потоа ги толкува овие сигнали и извлекува емоции од ова во реално време. Во комбинација со апликација за терапија и во соработка со терапевт, на овој начин може да се обучуваат емоционални и социјални вештини.

Зошто анализа на говор?

Говорните команди се користат во секојдневни ситуации. Затоа областа на интелигентна анализа на говор е суштински градежен блок за оваа терапија. Општо земено, начинот на кој некој кажува нешто дава многу повеќе индикации за емоционалната состојба на една личност отколку што се пренесува преку содржината на зборувањето.

Цел: база на податоци

Создавајќи ваков роботски систем, мора да се евидентираат податоци од споредбена група со типични деца, но и возрасни. Добиените модели на емоции потоа ќе обезбедат основа за споредба за акустична анализа со деца со аутизам. audEERING го води развојот на говорна база на податоци за целни групи – возрасни и деца. Затоа, испитаници од двете групи се поканети да снимаат дадени зборови и фрази за специфични емоции - на пример, фрустрација, радост, досада, интерес, тага, изненадување, гнев.

За поддршка и понатамошен развој на базата на податоци, AUDEERING создаде апликација за снимање како дел од истражувачки проект, што треба да го направи општото снимање на податоци поефективно.

Развивањето нови стандарди може да се гледа како тековен процес. Во моментов технологијата на audEERING од производот devAIce е имплементирана на ROS (Robot Operating System) (Робот оперативен систем). Технологијата засега работи со роботот Pepper. Проектниот партнер во Берлин, Универзитетот Hum-

boldt, проценува нова стратегија за интеракција користејќи препознавање на говорни емоции на audEERING во терапија со аутистични деца.

МОТИВАЦИЈА

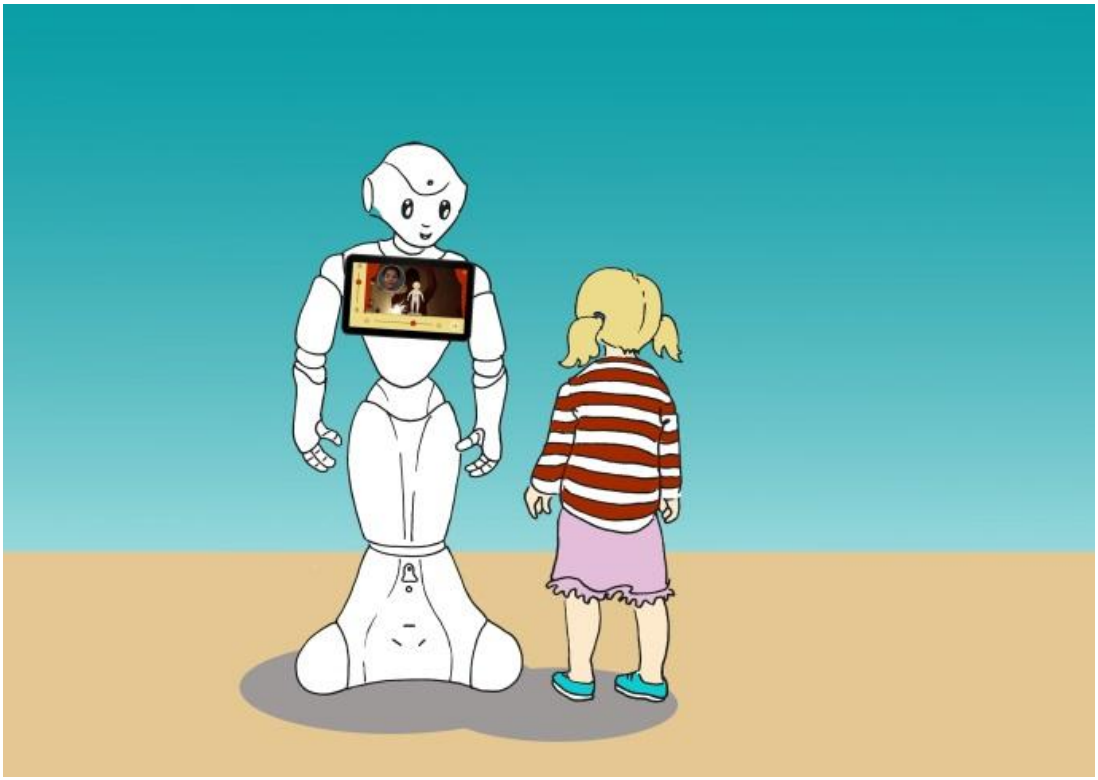
Разбирањето и изразувањето на социо-емоционални сигнали, како што се:

- израз на лице, и
- гласовна модулација,

е нарушен кај деца со аутизам. Додека на лица за човечка интеракција им е тешко да ги проценат, овие деца ги перцепираат роботите како попредвидливи и помалку сложени. Тие, исто така, често се технички заинтересирани и слободоумни.

ЦЕЛИ И ПРИСТАП

За да се развијат социо-емоционални комуникациски вештини на деца со аутизам, проектот ERIK развива и тестира нова форма на терапија со помош на роботски систем. Роботот “Pepper“ ги снима изразите на лицето и јазикот при интеракција со дете. Со играње со роботска топка “Leka“, пулсот може да се одреди и со помош на електроди. “Pepper“ ги толкува овие сигнали и извлекува емоции во реално време. Во комбинација со апликација за терапија “Zirkus Empathico“, може да се обучуваат емоционални и социјални вештини релевантни за секојдневниот живот. Со препознавање на интерес, фрустрација и здодевност на дете со аутизам, сценарија за терапија може индивидуално да се прилагодат. Користејќи гестови и движења на очите, “Pepper“ може да комуницира со деца на реален начин, намалувајќи ги стравовите кога имаат работа со луѓе.



Слика. Обука за социјални и емоционални вештини со робот

Роботи како дел од терапија: Интерактивни стратегии за деца со ограничени социо-емоционални вештини

Разбирањето и испраќањето социо-емоционални сигнали како изрази на лице и интонација се нарушени кај деца со аутизам. Користењето на роботи се покажа корисно во терапија на овие деца, од една страна затоа што роботите овие деца ги доживуваат како помалку сложени и затоа попредвидливи и помалку застрашувачки, а од друга страна затоа што децата со аутизам често имаат одлични интерес и многу силни страни спрема технички уреди. Овој факт се користи во истражувачкиот проект “ERIK“ за да се овозможи вистинска социо-емоционална интеракција помеѓу аутистични деца и роботи преку функции како што се:

- препознавање емоции и повратни информации во реално време, со што се создава иновативен терапевтски пристап.



Слика. Препознавање емоции и повратни информации во реално време во терапијата со аутизам

Цел: Развој на интерактивна роботска платформа за терапија

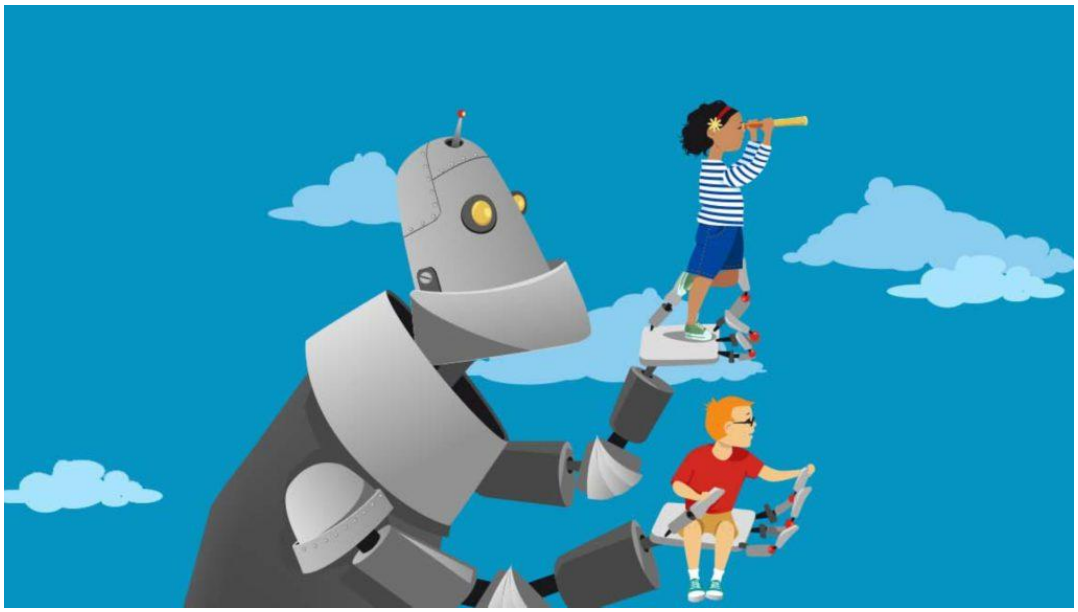
Во истражувачкиот проект “ERIK“ се развива роботска платформа која овозможува нови стратегии за интеракција во терапија на деца со ограничени социо-емоционални вештини. За разлика од претходни терапевтски пристапи, роботската платформа треба да може да ги препознава емоциите на детето и да генерира повратни информации во реално време врз основа на:

- изрази на лице,
- говор, и
- пулс на дете.

На овој начин роботскиот систем може да го поддржи детето во правилно препознавање и претставување на емоции на други и на тој начин да ги зголеми социо-емоционалните вештини на дете со аутизам. Благодарение на отворениот дизајн на интерфејси на роботска платформа, концептот на терапија може да се прилагоди и на потреби на:

- пациенти со деменција или пациенти со интелектуална попреченост,

со што се овозможува широк опсег на апликации.



Навремено оценување на цели на учење и стратегии за интеракција за функционален систем

Катедрата за производна автоматизација и систематика на производство (FAPS) на Friedriech-Alexander University Erlangen-Nuremberg (роботски систем), Fraunhofer IIS (физиолошко и видео-засновано препознавање емоции), audEERING GmbH (препознавање на емоции базирано на говор) и ASTRUM IT GmbH (евалуација и подготовка на податоци за интеракција), се вклучени во реализација на проектот. Во исто време, работна група за клиничка психологија и социјална интеракција на Универзитетот Humboldt ги дизајнира цели на учење и стратегии за интеракција на платформа за роботи и ја оценува платформата на роботи во тесна соработка со Fraunhofer Working Group for Supply Chain Services at Fraunhofer IIS со деца со аутизам, терапевти, лекари и роднини. Бидејќи евалуацијата со различни групи се одвива во текот на проектот, резултатите може директно да се влеат во понатамошниот развој на системот.

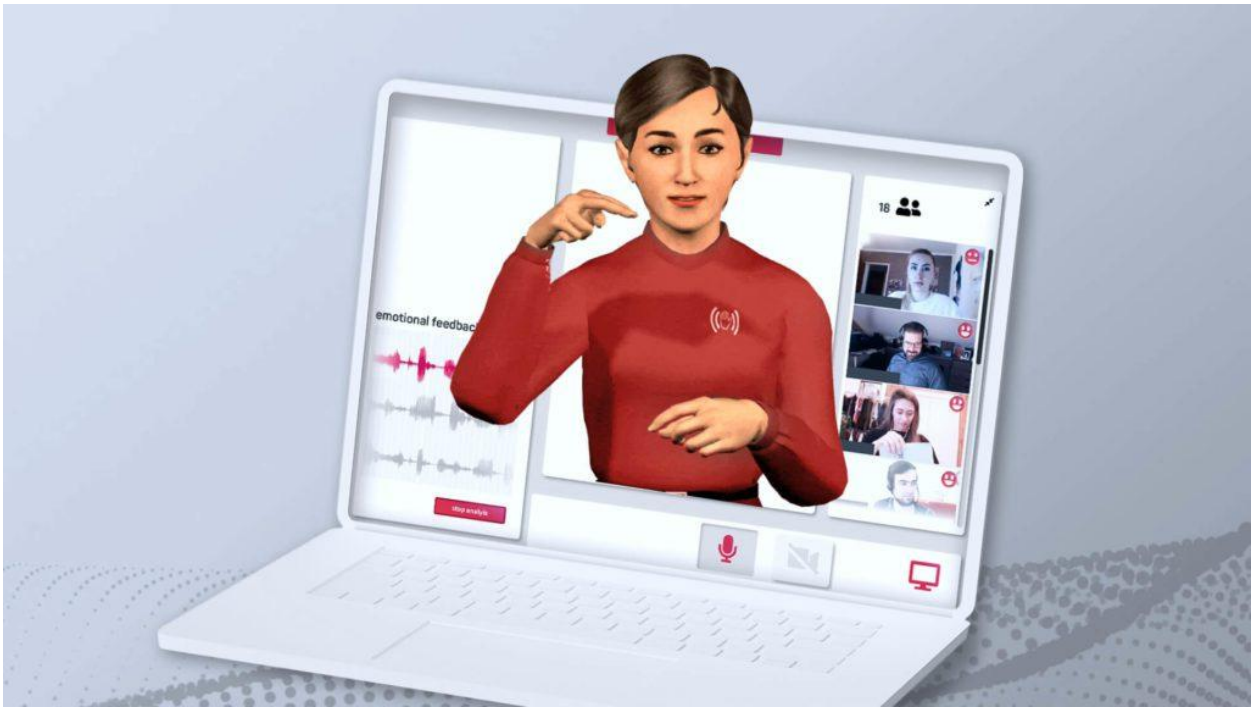
Во први чекор, се спроведени анализи на барања насочени кон корисникот и од техничка и од медицинска страна и се евалуирани етички, правни и социјални аспекти на системот на роботи. Со интегрирање на возрасни аутисти, родители на

деца со аутизам и терапевти во рана фаза во интервјуа и работилници, планираниот проект е во можност да се провери за неговата основна изводливост и прифаќање. Од друга страна, вреден придонес е добиен за специфичен дизајн на роботска платформа, која се базира на хуманоиден робот Pepper, и сценарија за интеракција. Овие сценарија за интеракција се развиени и интегрирани во системот на роботи заедно со јазично и физиолошко препознавање на емоции, како и интерфејс на терапевт кој им овозможува на терапевтите да ја контролираат и континуирано да ја следат интеракцијата робот-дете. Првите евалуации насочени кон корисникот на целиот систем, вклучително и во однос на неговото прифаќање и ефективност, веќе се извршени и продолжуваат да се оценуваат итеративно.

Роботска интеракција за деца од спектарот на аутизам

Како дел од студијата за истражувачкиот проект “ERIK“, HU Berlin создаде видео кое ја прикажува интеракцијата на роботите.

Вештачка интелигенција за анализа на говор – комуникација без бариери



Пред повеќе години образованието на глуви лица се фокусираше на учење да зборуваат, и се веруваше дека ова е единствен начин да се биде дел од општеството. За среќа, ова повеќе не е случај - но сè уште има недостаток на технологија за глуви лица.

Интелигентен автоматски превод на знаковен јазик

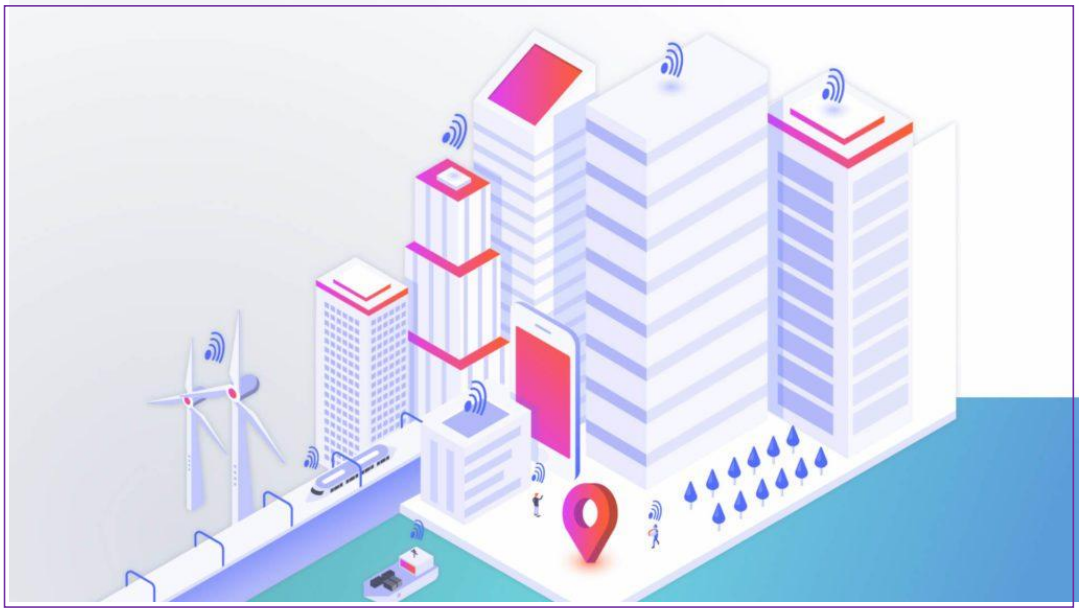
Истражувачкиот проект на ЕУ EASIER дизајнира, разви и потврди целосен повеќејазичен систем за машинско преведување кој ќе послужи како рамка за пристапна комуникација за глуви лица.

Аудио вештачка интелигенција за Паметни Градови

Пазарот за проекти за паметни градови е огромен и сè уште расте. На глобално ниво, 608 милијарди американски долари се потрошени на проекти за паметни градови во 2019 година. Прогнозите велат дека до 2024 година оваа сума ќе достигне повеќе од трилион американски долари ширум светот. Ова ја покажува важноста на овој сектор и е една од причините зошто AUDEERING учествува во истражувачки проект MARVEL финансиран од ЕУ.

Мониторинг за безбедност во реално време

Проектот MARVEL има за цел да ги поддржи работните текови на апликации и донесување одлуки во реално време управувани од податоци во современите градови. За таа цел, потребна е конвергенција на збир на технологии: ВИ, аналитика, мултимодална перцепција, софтверско инженерство и HPC како дел од парадигмата Edge-Fog-Cloud Computing Continuum, за поддршка на работни текови на апликации во реално време управувани од податоци и донесување одлуки во современи градови, покажувајќи потенцијал за многу ефикасно справување со општествени предизвици (кој оди подалеку од традиционалните Големи податоци). Со вакви технологии, општествените промени може да се решат многу ефикасно, од зголемување на јавната безбедност и заштита, до анализирање на сообраќајни текови и однесување во сообраќајот.



Како? Со справување со зголемени предизвици во синџирот на вредности на големи податоци во врска со стекнување (аквизиција) податоци, анализа и обработка на податоци, складирање и чување на податоци и визуелизација на податоци. Со приоритизирање и зајакнување на отворена наука и отворени податоци преку споделување корпус на податоци за поттикнување на истражувањ и иновации во мултимодална обработка и аналитика. Последно, но не и најмалку важно, со ангажирање на граѓаните да промовираат иновативни чекори.

Општествени предизвици. Една од целите на паметен град е настојување да се подобри квалитетот на животот на граѓаните преку инвестирање во ИТ решенија.

Сепак, додека користењето различни типови на ИТ решенија ја зголемува благосостојбата, сè уште има бројни општествени аспекти за кои луѓето генерално сé повеќе се загрижени, како што е:

- зголемување на контролата на градот,
- одбрана од терористички напади,
- криминални инциденти.
- природни катастрофи,
- управување со сообраќајот,

- фаворизирање економски развој итн.



Оттука, решенијата за паметни градови треба да се осмислат не само од техничка перспектива, туку и од општествена перспектива со цел да се разбере влијанието на технологијата во животите на луѓето.

Откривање акустична сцена за брзи одговори

Аудио-визуелниот мониторинг има значително влијание врз надзор за јавната безбедност и сигурност. Потенцијална безбедносна опасност е зголемена во преполни јавни површини како што се барови за пиење и места за рекреација и одмор, особено во текот на ноќта.

Врз основа на нови методи, пристапи и инженерски парадигми во мултимодално аудио-визуелно управување и обработка на податоци, MARVEL го прикажува потенцијалот за многу ефикасно справување со граѓанските предизвици.

За ова мултимодално следење на паметен град во реално време, audEERING го обезбедува своето добро познато откривање на акустична сцена засновано на распоредување на мултимодални сензори. Технологијата може да открие и идентификува настани во реално време и да го подобри времето на одговор на властите. Во случај на несреќа, на пример, полицијата и болничарите може да бидат предупредени многу побрзо и да се спасат животи.

Висококвалитетни микрофони и камери може да се распоредат во области во градот за хетерогени цели на сензори, стекнувајќи ситуациона свест за животната средина. Додавањето аудио за да се подобри видео надзорот може да го подобри безбедносното покривање надвор од покривање на камери за да помогне во откривањето и интерпретацијата на настаните додека се случуваат. Дополнително, следејќи го мобилниот пристап на толпата, мобилните уреди за паметни телефони, овластени со иновативни апликации, им овозможуваат на граѓаните да собираат аудио-визуелни текови на податоци и на тој начин да следат голем географски простор на градот без трошоци и без да вклучуваат специјализирани заробувачи на податоци. Мобилните уреди овозможуваат поврзување и на гео-референцирани текови на податоци на сензори со физички и социјални податоци. Следствено, граѓаните можат да стекнат поголемо знаење за нивниот урбан пејзаж, а градските менаџери можат да добијат подобро познавање за нивниот град, со што граѓанските политики поефикасно ги приспособуваат на реалните потреби на населението.

Со оглед на огромното количество аудио и визуелни податоци како дополнение на достапните Големи збирки на вистински и историски податоци понудени од хетерогени IoT сензори кои се расфрлани во опкружување на паметен град (на пример, безжични сензорски мрежи, безжични сензори кои мерат светлина, CO2, влажност, нивои на бучава), ќе се постигне мултимодална перцепција и интелигенција за аудио-визуелно препознавање сцени, откривање настани и свесност за ситуации без да ги прекршува етички ограничувања и ограничувања за приватност.

Општествени аспекти како приватност на податоци

Во Европа луѓето се загрижени за надзорот и приватноста на податоците, па затоа следењето на јавниот простор е тема што треба да се реши на транспарентен и отворен начин. Оттука, решенија за паметни градови треба да се замислат не само од техничка перспектива, туку и од општествена перспектива со цел да се разбере влијанието на технологијата во животите на луѓето. Конзорциумот за MARVEL има за цел да го обезбеди ова со учество на 17 организации од 12 земји.

Како луѓето чувствуваат симпатии преку гласот

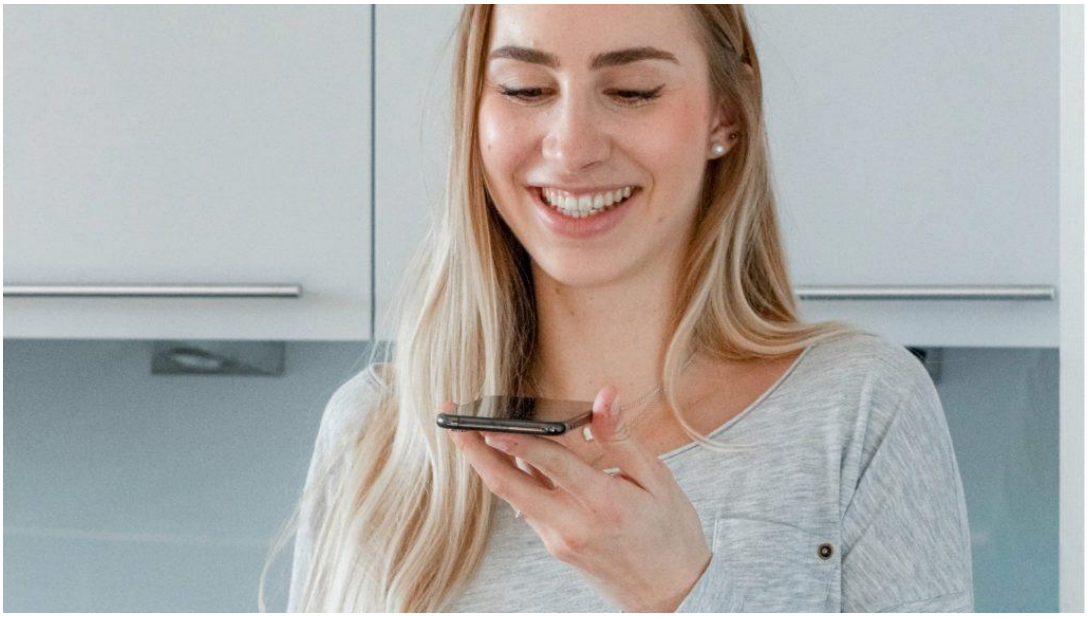
Допадливоста не е особина на говорникот, туку индивидуална проценка на примачот што креира важна особина (предуслов) за профитабилна говорна комуникација.

Социјален аспект на непознати гласови

Кога ќе слушнеме личност за прв пат, не можеме да се воздржиме од претпоставки, градење очекувања и да формираме прв став кон човечка или вештачка личност што ја пренесува тој глас. Се разбира, основата на која ги градиме нашите впечатоци е многу сиромашна, и се надеваме дека ги прилагодуваме нашите впечатоци со повалидни податоци од само овие неколку акустични (и визуелни) површни карактеристики. Сепак, овој прв впечаток може да има влијание врз нашето сопствено однесување, како што е поотворено или резервирано реагирање кон говорникот или како ги толкуваме изјавите на другиот. Може дури и целосно да избегнеме понатамошна интеракција, ако сметаме дека оваа личност е несакана.

Ова не важи само за човечки гласови во разговори лице-в-лице, туку и во телекомуникации, особено поради зголемен опсег и севкупен квалитет и природност на располагање. Во денешно време, гласови на лични асистенти, интерактивни гласовни одговори за самопослужување и паметни домашни уреди, достигнаа ниво на природност што понекогаш го отежнува препознавањето на нивната извештаченост. Како последица на тоа, социјалното прифаќање во првите средби станува се порелевантно не само во човечката комуникација, туку и во инте-

ракција човек-компјутер. Затоа, потребно е да се концентрираме на нејзините социјални импликации.



Аспекти на прв впечаток

Покрај лингвистичкото значење, гласот и стилот на зборување транспортира многу повеќе информации. Можеме, правилно или неточно, да заклучиме возраст, социјално и регионално потекло и пол. Можеме дури и да ја процениме привлечноста, расположението, ставот и аспектите на личноста. Сите овие фактори може да влијаат на тоа дали на слушателот или соговорникот му се допаѓа или не му се допаѓа говорникот. Особено позадината е склона кон позитивни или негативни стереотипи. И без никакво лично искуство, одреден стил на зборување, на пример на агент во центар за повици, може лесно (погрешно) да се припише на специфичен став или став наместо да се препознае индивидуална карактеристика.

Добро познати механизми на работа се преференции на луѓе со слично потекло и суб-култури, севкупно претпочитање на атрактивни гласови или претпочитање на стилови на зборување кое сигнализира интерес и пријателство, на пример со малку подигнат тон, живописни движења на тонот, малку зголемено темпо и звучна насмевка. Јасен глас, со негови спектрални импликации на помало

високофреквентно амортизирање и посилни хармоници, исто така, обично се фаворизира како јасен изговор. Такви механизми се прилично слични на визуелни знаци (движења на телото, облека, лице).

Проценка на допадливост на глас е деликатна тема

На работа, природно ги земаме предвид сите тие аспекти кога поединечно избираме, на пример, професионални гласовни таленти како говорници (презентери). Меѓутоа, автоматското проценување на допадливос на глас е деликатна тема. Еднаш, може да се предвидат само мнозински гласови за одредени домени. На пример, личност со силен регионален акцент. За разлика од проценка на пол или возраст, допадливоста навистина зависи од окото на гледачот. Исто така, конкретна ситуација ги дефинира основите за толкување дали одреден јасно изразен или нагласено пријателски стил е соодветен или не. Особено за синтетички гласови, несоодветноста на стилот со содржината и ситуацијата може да доминира во првиот впечаток, и на тој начин да замати јасен и привлечна порака, на пример.

Исто така, систематски ефекти пронајдени во истражувања, како и автоматски модели, се генерализираат од податоците. Затоа, треба да се претпочита чист глас на говорници кои покажуваат. Сепак, постојат луѓе кои имаат благ проблем со нивните гласни жици. Во зависност од точна манифестација, ова всушност може да резултира со тоа да биде сфатен како уште поинтересен, привлечен и допадлив.

Допадливоста не е особина на говорникот - тоа е перцепцијата направена од одредена личност

Затоа, допадливоста не треба да се смета како особина на говорник, како возраст или пол, туку како конкретен и специфичен, и како евалуација на (заклучени) состојби и особини на говорникот од страна на една индивидуа. Се разбира, постои договор меѓу слушателите кои ги оценуваат снимените примероци на говор. Ова може да се моделира и дава добра прва проценка за даден контекст. Но, овој договор не е толку силен како со основни емоции, на пример. Само земајќи ја предвид акустиката (т.е. квалитетот на сигналот, природноста на говорот, спектралниот приказ или интонацијата), секако ќе ја

ограничи точноста на одредување на допадливоста на гласот. И само со разбирање на специфики на содржината и социјалната ситуација, виртуелните личности ќе можат да станат допадливи по дизајн. Ова особено важи за социјални роботи и синтетички лица во виртуелни средини, кои исто така имаат визуелно или дури манифестирано тело за да одговара на нивниот глас и стил на зборување. Значи, ако гласот многу добро одговара на "битието", може да звучи малку искривено или невообичаено. И покрај овие предизвици, постои потенцијал и во индивидуална зависност помеѓу две лица во комуникација, бидејќи меѓусебната конверзија во стилот на говорот и изговорот се смета за позитивно за допадливост.

Многу начини да се биде допадлив врз основа на гласот

Постојат многу начини да се смета за допадлив врз основа на сопствениот глас. Со оглед на тоа што многу фактори влијаат, не постои можен општ совет за да се обезбеди допадливост. На пример, мажите не мора нужно да покажуваат низок глас за да бидат позитивно перцепирани. За дизајнери, ова е, се разбира, заеднички предизвик, а постојат методи на дизајнирање и корисничко истражување за да се најде решение за даден производ или услуга. За автоматски класификатори, наместо да се обидуваат да го покријат сложениот концепт на допадливост, насочување повеќе кон поврзани социјални сигнали на говорниците, како што се пријателство и ангажираност, може да биде поплодно за да се добијат точни предвидувања - барем за непознати гласови. Особено, како што се развиваат односите, овие гласовни карактеристики стануваат сé повеќе ирелевантни, бидејќи се користат само за идентификување на личноста - веќе ја знаеме неговата личност и ќе процениме дали оваа личност ни се допаѓа според начинот на зборување и гласот, или според новоста што ни ја соопштува.

Гласот е нова Крв

Крвта на човекот е основа за различни медицински апликации за дијагноза. Но, што ако и нашиот глас може да ни каже многу за моменталната состојба на здравјето на една личност? Дали гласот е нова крв? Со цел подобро да ги препознаеме акустичните карактеристики на обрасци на болест што се појавуваат

во човечкиот глас, прво мора да го разбереме механизмот на човечкиот производствен глас.

Вокален апарат

Различни структури во абдоменот, градите, грлото и главата мора да бидат координирани кога правиме звук: дијафрагмата, белите дробови, душникот и градниот кош служат како ветробран. Вокална поддршка, гркланот и глотисот (glottis) како генератор на вибрато и тон, фаринксот (pharynx), усната шуплина и носната шуплина како резонатор. Практично, целото наше тело влијае на звукот на гласот.

Разликување на болни од здрави лица според нивниот глас

Сето ова е предмет на процес на когнитивна контрола заедно со други области на нервниот систем. Кога когнитивниот или физичкиот напор, па дури и болеста, влијае на овој сложен процес, може да генерира карактеристики на јазикот што може да се користи за да се разликува болен од здрав.

Карактеристики се однесуваат на интонација, интензитет и темпо. Ако системи за вештачка интелигенција се обучени со соодветни податоци за звук и говор, кои обично се состојат од двојно аудио и придружна прибелешка, како што е дијагноза, тие можат да препознаат соодветни цели за учење.

Примери за физички и психолошки аспекти

Постојат два суштински аспекти кои можат да се забележат: **физички** и **психолошки** аспекти. ВИ може да ги “слуша“ сите овие аспекти со анализа на аудио сигнал и соодветно учење врз основа на многу податоци. Сензор потребен за снимање на гласовни податоци е микрофон, кој се наоѓа во скоро секој компјутерски уред денес како паметни телефони и таблети.

Пример кој особено се однесува на физички аспекти во однос на моторни нарушувања на производство на говор е Паркинсонов синдром. Кај ова невродегенеративно нарушување, меѓу друго, се оштетени мускулните групи вклучени во производство на говор. Така, личноста има проблеми со изговор на зборови, монотонија во однос на интензитет и интонација, намален акцент и груб квалитет на гласот.

За ментални аспекти и психолошка страна во однос на влијание на производство на звук и говор, препознавањето на емоции е совршен пример. Емоциите играат суштинска улога во обрасци на болести како што се **исцрпеност** и **депресија**, **аутизам** и **Алцхајмерова болест**. За препознавање емоции, гласот се смета за сè порелевантен индикатор.

Собирање податоци како база

Значаен фактор што влијае во анализа на говор базирана на вештачка интелигенција е, особено, присуството на соодветни гласовни податоци и квалитет. За да се подучат релевантните карактеристики на алгоритмот, потребен е голем број звучни и гласовни податоци. Потребни се и податоци за здрави лица со висока варијабилност, вклучително и разни други клинички слики и состојби и оние за болести.

КОВИД-19 како актуелен пример

Актуелен пример за користење на интелигентна анализа на говор и глас е откривање на COVID-19 од говор. Меѓу друго, вирусот го напаѓа и респираторниот систем, што се манифестира со следни симптоми: кашлање, стиснат глас, често болка во грлото и фаринксот, отежнато дишење или проблеми со белите дробови како што се шушкање на белите дробови.

Првичните студии покажуваат добри стапки на откривање на СОВИД-19. Сепак, комбинација на анализа од говорни и звучни настани е најдобра за да се даде уште повалидна изјава за симптоматологија на COVID-19 и да се користи како алатка за дијагностицирање. Ова може да даде значаен придонес во ограничување на пандемијата, особено во прекинувањето на синџири на инфекција и откривањетна болеста во рана фаза.

Емоционална интеракција – алка (Link) што недостацува

Емоционална интеракција

Следен чекор е да се додаде уште еден слој со негови димензии и режими на нашата интеракција. Ако некогаш сме патувале во странска земја, каде што не можеме да го зборуваме јазикот на жителите, сигурно знаеме што значи да разговараме со луѓето и да ги слушаме без да разбереме ниту еден збор што го

кажуваат. Сепак, сè уште зборуваме на нашиот мајчин јазик и го користиме нашиот тон како алатка заедно со говорот на телото за да се повземе и да комуницираме. Тоа е магијата на емоционална интеракција.

Препознавање на емоционални состојби преку гласот

Не мора да сме светски патник за да знаеме дека ние како луѓе се поврзуваме преку древна форма на јазик: **јазикот на емоциите**. Иако степенот на нашите изрази е под силно влијание на средината во која сме израснати и како се мисли дека ги регулираме, основен образец за главните емоции сè уште е ист кај сите култури. Нашата смеа и плач, нашиот гнев и возбуда се изразени со ист вокален тракт сличен на древниот хомо сапиенс пред 60.000 години. Генерација по генерација, ние сме обучени да ги детектираме овие индиции во гласот и веднаш да ги поврзуваме со емоционални состојби или модели на однесување.

Илјадници гласовни примероци за секоја емоција

Дали и вештачката интелигенција може да го научи тоа? Користејќи методи за **длабоко учење** (deep learning), компјутерите се хранат со илјадници и илјадници гласовни примероци за секоја емоција. Ова е ист процес што се случува во нашиот мозок уште од детството. Потребна е голема компјутерска моќ за да се обучи моделот, но штом ќе заврши, може да работи во реално време на кој било уред. Во денешно време, вештачката интелигенција може да открие општи емоции толку добро колку што можат луѓето. Апликации на оваа иновативна технологија се безбројни.

Емоционална вештачка интелигенција (Emotion AI)

Emotion AI ни овозможува да го завршиме циклусот на емоционална интеракција. Тоа значи дека компјутерите можат да откријат и да бидат свесни за емоционална состојба на корисникот и соодветно да реагираат.

- **Афективно пресметување** (Affective computing) е проучување и развој на системи и уреди кои можат да препознаат, интерпретираат, обработуваат и симулираат човечки афекти.

Тоа е интердисциплинарно поле кое опфаќа компјутерски науки, психологија и когнитивна наука.

Идејата за завршување на овој циклус не е работа од 2020 година. Со секој напредок во технологијата и откривање на сензори, научниците се обидуваа да го завршат циклусот. Користење на отчукување на срце, пот на нашите раце, движења на очите и мозочни бранови се некои примери за овие напори. Иако сите тие се покажаа како ефикасни во лабораториски услови и истражувања во мали размери, тие не ја променија индустријата за видео игри. Причина за тоа е едноставна: **скалирање** и **достапност**.

Аудио вештачка интелигенција

И покрај придобивките од истражувањата, природната интеракција и дополнителните сензорски системи, видео гејмерите не можеа и не сакаа да прикачат ЕКГ (Електро-кардиограм) или GSR (Галвански одговор на кожа) пред да го притиснат копчето за репродукција. Така, идејата за интеграција на афективни компјутери и видео игрите остана во научни трудови и никогаш не стигна до поширокатa јавност.

Но, што ако можеме да користиме сензори кои веќе се таму и да користиме начин на афективно пресметување што може да се скалира? Ова прашање нè доведува до Audio Emotion AI. Микрофоните се насекаде и зборувањето е најприроден начин на интеракција за нас. Користејќи Audio Emotion AI, можеме да го завршиме циклусот на интеракција. Овој пат сензорите се таму на кој било уред, моделите се со мала тежина и афективниот режим на интеракција е природен.

Придобивки од интегрирање на емоции

Со отстранети сите бариери, прашањето е зошто всушност треба да го направиме тоа? Која е користа од интегрирање на емоции во нашите секојдневни интеракции? И која е опасноста од недостаток на емоции во тие интеракции?

Нови начини на интеракција човек-машина во видео игри

Кога пишуваме СМС-порака на наш пријател, му се јавуваме или го сретнуваме во кафуле, ние сме во интеракција со него. Секој разговор може да биде интеракција. Интеракција се случува кога два или повеќе предмети или агенти влијаат еден на друг. Пред неколку векови, луѓето главно комуницираа меѓу себе и нивните домашни животни, а благодарение на еволуцијата, ние сме опремени за тоа.

Со доаѓањето на машините во нашиот секојдневен живот, повеќе дисциплини како човечки фактори, комуникациски дизајн, информатичка технологија итн. се овозможи да се осигураме дека можеме да комуницираме со нови објекти (интеракција човек-машина) и ентитети во нашата средина, толку природно како што може овчарот со своето куче.

Аудио излез

Аудио излезот помина низ долг процес. Се започна со single звучници кои можеа да произведат некои поединечни тонови, а подоцна се воведоа звучници што можеа да ги произведат сите звуци. Скокот беше земање предвид на 3D просторот и обработка на аудио во повеќе канали. Исти процес се случи со слушалките бидејќи тие станаа подобри во поддршка на 3D аудио и изолација. Во денешно време, можене да набавиме 3D слушалки со активно поништување на шум што се прилагодува на околината за подобри перформанси како производ за широка потрошувачка, на пример, од Jabra. Следен скок е **хаптички звук** (haptic sound). Компании како **Woojer** веќе создадоа елек и ремен за да ни помогнат да го почувствуваме басот.

Подобрени работни услови за постари работници

Околу 20% од работоспособно население во Европа е постаро од 55 години. Европското население старее и добри работни услови се поважни од било кога. Во општество каде вработувањето е продолжено на повисоки возрасти, проект како WorkingAge има за цел да промовира здрави навики во работни средини фокусирајќи се на луѓе на возраст над 45 години. audEERING е дел од конзорциум заедно со Кембриџ Универзитетот во и многу други универзитети и компании.

Истражување за работни услови во канцеларија, далечинско работење и производство

Цел на WorkingAge е интегрирано дигитално решение. Со проучување на три типа работни средини (канцеларија, далечинско работење и производство) овие средини се погодни за постари работници. WorkingAge има за цел:

- Намалено присуство на стрес.
- Подобрени работни перформанси поради поголема мотивација и вклученост во одлучувањето.
- Поголема доверба во можност за промена или наоѓање работа.
- Зголемување на чувство на припадност.
- Зголемување на самодоверба.
- Зголемување на интеграција на работници во работна средина.

Алатки за следење на нивои на стрес

За таа цел, проектот создаде алатка наречена WorkingAge of Wellbeing (WAOW) алатка за да обезбеди помош во секојдневни рутини како препораки, избегнување ризици и потсетници. Однесување, здравствени податоци и преференции на работникот се следат за пријателски интеракции со корисникот. audEERING обезбедува следење на бучава во опкружувањето (ниво, гласовна активност) и препознавање на емоции од гласот. Со следење на овие фактори, може да се открие стресот и да се прилагодат работните услови на начин што најдобро одговара на потребите на работниците.

Скалабилен производ за компании

Оваа алатка WAOW ќе еволуира во скалабилен производ за компаниите. Бидејќи најниската стапка на вработеност во Европа е пријавена за група на возраст од 55 до 64 години, проектот има за цел да го задржи вработувањето на работниците од оваа возрасна категорија, преку обезбедување алтернативни работни шеми. Компаниите можат да ја користат алатката WAOW за да ги поттикнат своите вработени да го подобрат својот професионален живот според потреби што ги имаат во нивната специфична фаза од животот.

ЗВУЧНИК: Аудио вештачка интелигенција за паметни звучници (Smart Speakers)

Паметни звучници денес се дел од дневните соби на многу луѓе, но досега на пазарот има само уреди од САД: Alexa, Google Home, Apple HomePod, и други. Истражувачкиот проект SPEAKER има за цел да го промени тоа. Целта на овој проект е да се развие европски паметен звучник за да се натпреварува со американските производи на пазарот. audEERING е дел од конзорциум што го развива ова решение.

Европска безбедност на податоци во паметен звучник

Па, зошто ни е потребно европско решение, ако американските производи во моментов се најсовремени? Одговорот е лесен: не е можно да се имплементираат европските стандарди за безбедност на податоци со решенија од странство. Понатаму, Европа не сака големиот потенцијал на овој пазар да го остави само на американски компании.

Асистент прилагоден на потреби на корисникот

audEERING ја обезбедува својата добро позната аудио вештачка интелигенција за да ги подобри функционалностите на европскиот производ. Обезбедува неколку основни функционалности:

- откривање на особини на звучници,
- откривање на гласовна активност (VAD- Voice Activity Detection),
- детекција на пол, и
- анализа на акустична сцена.

Откривањето на карактеристиките на говорникот и полот има за цел приспособени предлози од паметниот звучник. Тоа на пример, да биде можно да се открие дека корисникот се чувствува под стрес од напорниот работен ден кога влегува во собата. Асистентот потоа може да ја предложи неговата омилена музичка листа за релаксација.

Анализа на акустична сцена и откривање на гласовна активност (VAD) се користи за да ги филтрира звуци во просторија.

Од истражување во пракса

За да се направи оваа визија реалност, создаден е сеопфатен екосистем на компании. Голема индустрија, средни компании, стартапи и истражувачки

партнери обезбедуваат иновации и трансфер од истражување во пракса. Два Fraunhofer институти го водат конзорциумот и големи играчи како SAP, Deutsche Bahn, Siemens и Bosch се дел од него. Исто така и audEERING е дел од овој конзорциум.



Безбедност на податоци кога се користи вештачка интелигенција во било каков производ

Безбедност на податоци е клучна кога станува збор за користење на вештачка интелигенција во секаков вид производ. Кои се најважни аспекти за безбедност на податоци и како да се постапува со чувствителни податоци.

1. **Што се подразбира под поим безбедност на податоци**?

Безбедноста на податоците, има неколку аспекти: **внатрешно ракување со податоци** (internal handling of data), на пример, за човечки ресурси и **ракување со надворешни податоци** (handling of external data), на пример, од клиент. Кога се работи со податоци за клиенти, има неколку правила:

- **Само внатрешно складирање податоци** (Internal data storage only): Податоците се зачувани на **внатрешни сервери** (кај истражувачки компании) и пристапот е ограничен. Нема складирање податоци во **облак** или на лаптопи на вработени.

- **Пристап само за овластени лица** (Access for authorised persons only): **Прво**, процесот на доставување на податоци треба да биде безбеден. Од испорака до завршување на еден проект, јасно е кој има пристап и му е дозволено да работи со податоците. Понекогаш, има клиенти кои даваат само ограничени права за пристап, па мора да се постават одредени правила на серверите. Потоа, треба да се следи и кој и кога работел со одредени податоци.

2. Како се обезбедува заштита на податоци во audEERING?

Лице задолжено за заштита на податоци го пренесува своето знаење на истражувачкиот тим. За снимање на податоци од испитаници, се обезбедува:

- креирање политики за снимање податоци за специјални апликации и нивно предавање на лице задолжено за приватност на податоци, и
- кога се складираат (меморираат) податоци, мора да се проверат лиценци за сите информации кои се користат.

3. Како се обезбедува заштита на податоци кога се снимаат чувствителни гласовни и здравствени податоци?

Податоците за говорот се секогаш посебни затоа што не може целосно да се анонимизираат податоците. **Звукот на гласот** на една личност е толку уникатен, што дури и без да се даде име или други лични податоци, некој потенцијално би можел да ја препознае оваа личност.

Кога станува збор за порталот AI SoundLab, не се снимаат само гласовни податоци, туку се снимаат и мета податоци. Колку податоците стануваат почувствителни, толку се повисоки стандардите, кои мора да се чуваат во однос на заштита на податоци и толку е потешко да се складираат податоците. Ова значи дека треба да се обезбеди овие податоци да се евидентираат, испорачуваат и чуваат многу строго во согласност со европските закони за заштита на податоци.

4. Дали audEERING има право да ги користи собраните информации за други производи?

Ова навистина зависи од видот на податоци што се добиваат и од лиценцата што се применува. Ако се добиваат податоци од клиенти, во некои случаи користењето е многу ограничено. За здравствени податоци, на пример, може да се случи, воопшто да има само еден или двајца вработени на audEERING, на кои им е дозволено да ги обработуваат овие податоци. Користењето потоа е исто така ограничена на овој специјален проект со клиентот. Понатаму, мора да се избришат податоците потоа за да не може да се користат податоците за ништо друго.

Во други случаи, клиентите се во ред да споделуваат податоци со audEERING и им дозволуваат да ги користиат карактеристиките на аудио-датотеките за градење модел. На пример, ако се добијат податоци од универ-зитет, постојат некои отворени лиценци кои овозможуваат да се користат овие податоци за сите различни видови производи, истражување и градење модели.

5. Која е разлика помеѓу анонимно и псевдонимно складирани податоци?

Аудио податоците се единствени (unique). Ако ја познаваме личноста што зборува, би ја препознале. Ова значи дека аудио податоците не можат целосно да се анонимизираат, доколку се работи за вистински необработени аудио податоци, кои не се обработени, на пример, со посебен филтер.

Во audEERING ги псевдонимизираат аудио податоците, т.е. се собираат мета податоци поврзани со необработените аудио датотеки. Еве еден пример како функционира ова: Ако се бара само пол или возраст, може да има неколку лица во базата на податоци кои имаат ист пол и возраст и никој навистина не би знаел која е оваа личност. Меѓутоа, ако се бара и националност, податоците добиваат помалку псевдоним и во одреден момент веќе не се ни псевдонимни, бидејќи од сите овие податоци може да се извлече која е таа личност. Значи, затоа треба да соберат што е можно помалку податоци, но колку што е потребно за да се осигура дека тие ќе добијат што е можно повеќе псевдоним.

6. Кои се специфики околу безбедност на податоци во здравствениот сектор (општо)?

За здравствени податоци законите се многу построги отколку за други податоци. Доколку некоја компанија сака да создаде здравствен производ врз основа на вакви податоци, потребни се одредени сертификати. Значи, ако една компанија сака да влезе во здравствениот сектор, има многу повеќе бариери. За идни медицински апликации, се работи со партнери бидејќи audEERING не се медицинска компанија. На пример, audEERING работи со неколку болници на проекти.

Апликации, типови ВИ и ефекти од ВИ

Апликации на вештачка интелигенција

Постојат два типа на работни места за вештачка интелигенција:

1. Работи што ни го олеснуваат животот.
2. Нови работни места што ние како луѓе не можеме да ги работиме.

На пример, апликацијата за препознавање емоции во телефонски (call) центар може да им помогне на агентите објективно да размислуваат за нивниот тон и да се приспособат во реално време доколку е потребно. Освен тоа, тие исто така можат да видат како се менува тонот на клиентот во текот на разговорот и колку среќа тие носат на овој свет. Овој пример јасно покажува примена на ВИ во која им помага на луѓето рамо до рамо и ги зголемува.

Да ги земеме гласовните биомаркери како друг пример. Науката покажува дека некои болести прават промени во нашите гласни жици и нашиот глас. Овие промени се толку суптилни што не можеме да ги слушнеме, меѓутоа, вештачката интелигенција може да ги разгледа **промените во фреквенциите** и **аудио брановите** и таму да најде обрасци кои инаку ни се скриени. Потоа ќе ни помогне да откриеме болести како Паркинсон и други. Можеме да го дониграме нашиот глас во **AI-sound-lab** и да имаме улога во создавање на **здравствена вештачка интелигенција** (health AI).

Ако ја погледнеме индустријата за игри, таму можеме да ги видиме и двете од овие категории. Забавно набљудување може да им помогне на компаниите во

процесот на тест за играње да се осигураат дека нивната игра резонира со нивната публика. Во овој пример, вештачката интелигенција ни заштедува многу време за да ја слушаме секоја сесија и ни кажува на кои делови е поважно да се фокусираме. Од друга страна, **entertAIn play** им помага на дизајнерите и развивачите на игри да креираат игри и да имплементираат нови емоционални интеракции кои инаку не се можни. Во овој случај, вештачката интелигенција овозможува да имплементираат сосема нови сценарија.

Типови на вештачка интелигенција

Денес кога знаеме повеќе за вештачката интелигенција, веројатно се прашуваме за иднината на вештачката интелигенција и какви различни типови ќе се појават во иднина.

Примена на вештачка интелигенција за препознавање емоции со помош на аудио е тесен случај на користење. Затоа се нарекува **тесна Вештачка интелигенција** (ANI - Artificial Narrow Intelligence). Може да извршува задачи како и човеч или уште подобро.

Повеќе научници мислат дека до крајот на овој век ќе имаме вештачка интелигенција која хипотетички ќе може да научи и да ја разбере секоја интелектуална задача што ја може човечкото суштество. Ние обично ја нарекуваме **општа вештачка интелигенција** (AGI - Artificial General Intelligence). Тоа значи дека може да имаме агент за вештачка интелигенција на работа што може да ни помогне на похуман начин отколку со часови да пребаруваме на Google за да најдеме одговор.

Хипотетички кажано, научниците веруваат дека во одреден момент во иднината на човештвото, ќе имаме **супер вештачка интелигенција** (ASI - Artificial Super Intelligence). ASI е **хипотетички агент** кој поседува интелигенција која далеку ја надминува онаа на најпаметни и најнадарени човечки умови. Со други зборови, односот на ASI со нашата интелигенција ќе биде односот на користење на нашите мускули за работа со лопата и користење на багер со тркала со корпа (BWE - bucket-wheel excavator). На машинските инженери им требаа неколку децении да преминат од парна машина до BWE, а сега повторно сме сведоци на

тоа на софтверските инженери и истражувачите на вештачка интелигенција во нивната област.

Ефекти на вештачката интелигенција

Кога ја гледаме еволуцијата на вештачката интелигенција и промените во дефиницијата, наидуваме на концепт наречен “**Ефект на вештачка интелигенција**“ (“The AI Effect”). Со едноставни зборови, тоа значи што и да можат компјутерите да направат и ние го разбираме, повеќе не го нарекуваме ВИ. Во минатото, вештачка интелигенција беа некои алгоритми кои можеа да ја разберат животната средина и да комуницираат со неа. Сега, тоа е систем кој може да учи и извршува задачи кои се сметаат за остварливи само од луѓето, а можеби во иднина нашите очекувања од вештачката интелигенција би биле да извршува задачи над човечки и тоа би било сосема нормално за нас.

Како треба да звучи ВИ кога машината зборува со нас?

Како треба да звучи машината? Опкружени сме со машини што зборуваат, но дали нивниот глас е соодветен?

Интеракција со роботи станува сè повеќе дел од нашиот секојдневен живот, без разлика на обликот и големината или дали имаат тело пред сè. Домашна автоматизација, паметни звучници, апликации за разговор, роботи за грижа: сите овие користат најприроден комуникациски интерфејс што го измисли човекот: говор, но повеќе од нив имаат стандардни гласови кои не се поврзани со телото што го поседуваат.

Друг аспект на роботско вокално изразување е стилот на зборување кој во идеален случај би се прилагодил на комуникациската ситуација, но барем треба да биде соодветна на задачата за која е дизајниран роботот. Ако на пример роботот е дизајниран како играчка за интеракција со деца, стилот на гласот на човек-сидро веројатно не е добар избор.

Дали роботите треба да звучат “природно“?

Како што покажуваат повеќе истражувања, гласовите на роботите не треба да бидат колку што е можно повеќе слични на луѓето. Дизајнирање на глас повеќе

вештачки може да биде дури и решение за да се избегне **ефект на чудна долина** (the uncanny valley effect) кој вели дека ако вештачко суштество звучи природно како човек, но не сосема, тоа не им помага на корисниците да се чувствуваат удобно, но **може да изгледа морничаво**.

Понатаму, гласот сличен на човек за робот кој тогаш не прави ништо за да ни донесе малку храна може да доведе до искушение да разговараме за нашите брачни проблеми, што веројатно ќе резултира со фрустрација.

Во обемни истражувања, роботски гласови од филмови и игри беа испитани и опишани според акустични параметри. Повеќе од нив се одликуваа со доброволен висок степен на извештаченост што „... може да се постигне со мало зголемување на тонот, проследено со додавање хармонии и воведување одредено ехо“. Се разбира, фактот дека луѓето беа изложени на одредени карактеристики на роботски (и вонземски) гласови во медиумите уште од 50-тите, води до одредено очекување.

Така, иако според истражувачки сознанија ниту еден синтисајзер за говор никогаш не зборувал со рамна контура, штом почнеме да зборуваме на монотон начин луѓето нé мешаат со робот.

Гласови на големи и мали роботи

Кога роботите зборуваат, гласот обично не се генерира од механички вокални органи, туку со дигитална синтеза на говор, така што совпаѓањето помеѓу гласот и изгледот не доаѓа природно. Сега постојат некои основни закони на физиката кои го поврзуваат звукот на гласот со телото што го произведува, на пример, подолги вокални акорди се нишаат побавно. Ако совпаѓањето е лошо, корисниците можеби нема ни да сфатат дека гласот потекнува од робот.

Друго интересно прашање: дали е поубаво да се има една глобална интелигенција позади која ни зборува преку мноштво машини или би сакале секоја машина да има свој глас и личност? И ако е така, која личност и што е со прашањето за полот?

Дали роботот треба да звучи среќно?

Соодветноста на емоционален синтисајзер секако зависи првенствено од неговата примена: синтисајзер што дава глас на цртани фигури исполнува различни барања отколку систем да креира глас на лице со посебни потреби. Забавни, на пример емотивни честитки, аватари за разговор, игри, веродостојни ликови, адаптиран дизајн на дијалози, адаптиран дизајн на личности, рекламирање специфично за целна група, веродостојни агенти, вештачки луѓе, се случаи на користење за емотивен говор, каде што апликациите се тесно поврзани со развојот на вештачката интелигенција. Бидејќи емоциите и интелигенцијата се тесно измешани, потребна е голема грижа кога компјутерските системи се чини дека реагираат емоционално без интелигенција да ги исполнат очекувањата на корисникот во однос на способности за дијалог.

7. Емоционални аватари и решение за гласовна вештачка интелигенција

Аватар во хиндуизмот се однесува на манифестација на божество или ослободена душа во телесна форма на Земјата. Во денешно време, луѓето почесто го поврзуваат со икона или фигура што претставува одредена личност во видео игра, интернет форум, социјални медиуми итн. Ние користиме аватари за да го покажеме нашиот идентитет или да претпоставиме други идентитети, а кога станува збор за идентитет, сакаме да се осигураме дека се изразуваме како што сакаме.

Пазарот ја знае вредноста на самоизразувањето и брзо капитализира од тоа. Овој бизнис модел наликува на живот и несвесно луѓето подобро го разбираат. Играњето (живот) е бесплатно, но плаќаме за дополнителни предмети како што е брендирана облека за подобро да се изразиме.

Пример, играта Fortnite е бесплатна игра, но заработи приближно 70% од својот годишен приход од 5,1 милијарди долари преку продажба на брендирана облека (T-shirt маици). На пример, Epic заработи 50 милиони долари CAMO од NFL сет за брендирани маици, а не беше ни најпродаван.

Следни чекори во Metaverse

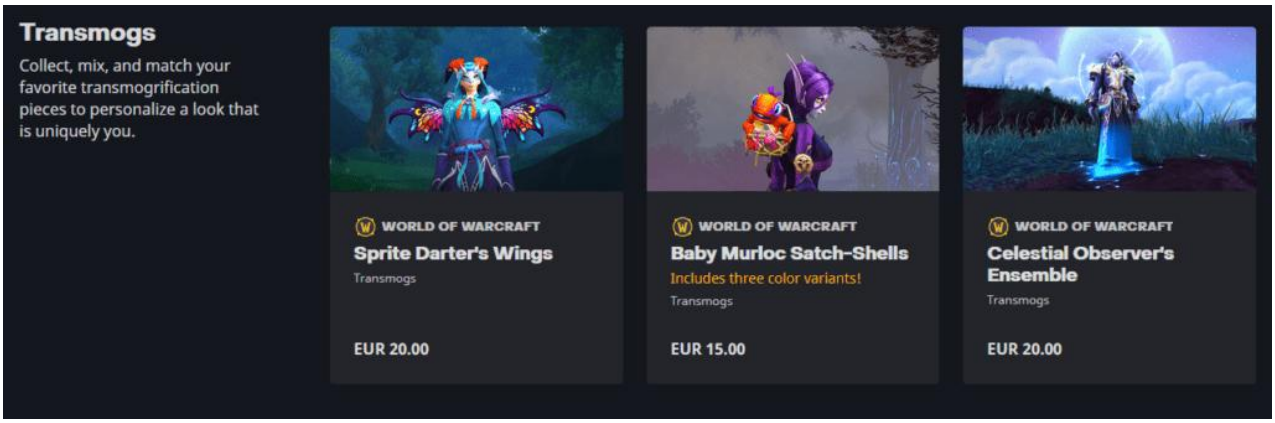
Со доаѓањето на Metaverse и преминувањето од 2D на 3D искуства, потребата за подобри аватари е поизразена. И оние кои даваат најдобри решенија и медиуми победуваат во оваа рунда. Зошто луѓето плаќаат милијарди долари за овие аватари и кој може да биде следниот чекор?

Речиси секој F2P (бесплатно за играње) како Fortnite и Dota генерира најголем дел од своите приходи преку продажба на скинови (skins). Дури и игри како World of Warcraft исто така нудат скинови што не се достапни во играта. И како тие го продаваат тоа:

- Секој може да го персонализира изгледот на своите омилени ликови во видео играта како што тој сака, т.е ја креира облеката (избира свои омилени парчиња облека) - transmogrification (avatar skin).



Звучи како да имам свој плакар, за да можам да го креирам своето дигитално јас (digital self). Значи, веќе сме на ниво каде што можеме да купиме облека за наше дигитално претставување, а слободата на дигиталниот свет ни дозволува да бидеме што било. Во оваа рунда можам да играм како Batman, а во следната рунда како Harley Quinn.



Опсег на аватари - од експресивен до фотореалистичен

Прво, неопходно е да се направи чекор назад и да се види спектарот на аватари. На едниот крај од овој спектар, сакаме да бидеме токму свои, па има смисла да се насочиме кон фотореализмот. На другиот крај на овој спектар, сакаме да бидете нешто што најдобро нé изразува.

Луѓе кои не се од овој домен можеби посакуваат да се претпоставени целосно како себеси, и тоа е најдобар начин да бидат експресивни како себе.. Сепак, тие забораваат дека луѓето одат во дигитален свет “токму“ поради таа причина - да бидат надвор од нивниот физички изглед и да искусат работи што се невозможни во физичкиот свет.

Најдобро претставување за секоја ситуација

Очигледно, постојат различни случаи на користење кои фаворизираат различни точки на овој спектар. На пример, ако создаваме VR искуство за обука и терапија за фобија, веројатно е дека фотореализмот ќе им обезбеди на корисниците подобри резултати. Меѓутоа, ако правиме VR игра како Fortnite, тогаш поголема е веројатноста експресивноста на аватарите да ја направат нашата игра успешна.

Фокусирајќи се на “експресивната“ страна на спектарот, прашањето е: Дали нашиот аватар може да се промени врз основа на нашиот емоционален израз?

Ако е така, како? И дали е корисно?

Баскиско истражување

Basque Research, исто така, спроведе студија наречена Аватари како комуникатори на емоции. Во оваа д-р. теза, проф. Амалија Ортиз предлага дека аватарите се еден од најдобрите начини за компјутерските системи да емитуваат невербални информации. Амалија создаде VR систем во кој корисниците доживуваат патување низ историјата. Оваа апликација, наменета за е-учење ги покажа следните резултати:

- корисниците претпочитаат интеракција користејќи аватари бидејќи сметаат дека е пријатна, кориснички ориентирана и забавна.
- корисниците се способни да препознаат повеќе емоции или изрази на лице и веруваат дека информациите ќе бидат подобро објаснети ако ги дава емотивен аватар.

Универзитетот во Копенхаген

На Универзитетот во Копенхаген, Aske Mottelson и Kasper Hornbaek од Катедри за психологија и компјутерски науки ги здружија силите за да ја проучат интеракцијата помеѓу афектот и сопственоста на виртуелно тело за емотивни аватари. Резултатите од оваа студија покажуваат дека некои манипулации со аватари можат суптилно да влијаат на афектот. Манипулација со лице се покажа како најефикасна во овој поглед, особено за позитивен афект. Исто така, сопственоста на телото покажа умерено влијание врз афектот. Резултатите покажуваат дека сопственоста на телото варира со валентноста, но не и со возбуда, и варира со позитивно, но не и со негативно влијание.

Нашиот мозок реагира на емоции во аватари

Прашање што може да си го поставиме е “зошто треба да се грижиме за емоциите во аватарите?“. Афективни невролози и социјално-когнитивни истражувачи беа активни за да ни помогнат да одговориме на ова прашање:

Како нашиот мозок реагира на емоции во аватари за разлика од луѓе?

Студија тврди дека страшни изрази во човечки и аватарски лица предизвикува слични одговори во амигдалата. Амигдалата е дел од мозокот одговорен

за емоции. Ова може да објасни зошто сме толку многу трогнати од она што се случува со ликовите во видео игрите.

Општественото влијание на игрите

Оваа пандемија нè научи дека работите што ги земаме здраво за готово може да ни се одземат за кратко време. Недостаток на социјални интеракции втурна многу луѓе во долина на депресија за време на оваа пандемија. Сепак, докази покажуваат дека играњето видео игри може да ја олесни осаменоста за време на пандемија на коронавирус. Зошто? Бидејќи што се однесува до нашиот мозок и амигдала, луѓето и нивните аватари не се толку различни.

Така, колку поприродни се овие интеракции и аватари, толку попријатни можат да бидат овие искуства. Со доаѓањето на Metaverse и ветувањето дека ќе поминуваме повеќе време во виртуелни светови, треба да се погрижиме таквите интеракции да бидат колку што е можно подобрени. Ваквите подобрувања можат да ни помогнат подобро да ја изразиме нашата индивидуалност и во овие изрази да ги надминеме границите на нашиот физички свет.

Гласовна вештачка интелигенција како клучен фактор

Клучен фактор во изразувањето се емоциите. Без да ги откриеме емоциите на играчите, немаме начин правилно да го промениме нивниот аватар за да го демонстрираме нивниот израз и да ја покажеме нивната индивидуалност.

Користејќи entertAIn play, секој развивач на VR и Metaverse може да го користи микрофонот на која било VR слушалка за да ги открие емоциите на своите играчи. Работи беспрекорно на уред без потреба од снимање и испраќање на аудио податоци на надворешен сервер, со што се зачувува приватноста на играчите.

Voice AI е алката што недостасува во овој долг синџир на изразување во Metaverse, и оваа технологија е решение за оваа употреба. Се нудат мноштво модели со вештачка интелигенција кои можат да се користат надвор од кутија (out-of-the-box) или да се калибрираат со најголема прецизност.

devAIce® XR ни овозможува да ја вклучиме Voice AI во било кој свет

Voice AI во **XR Humanizing Virtuality**

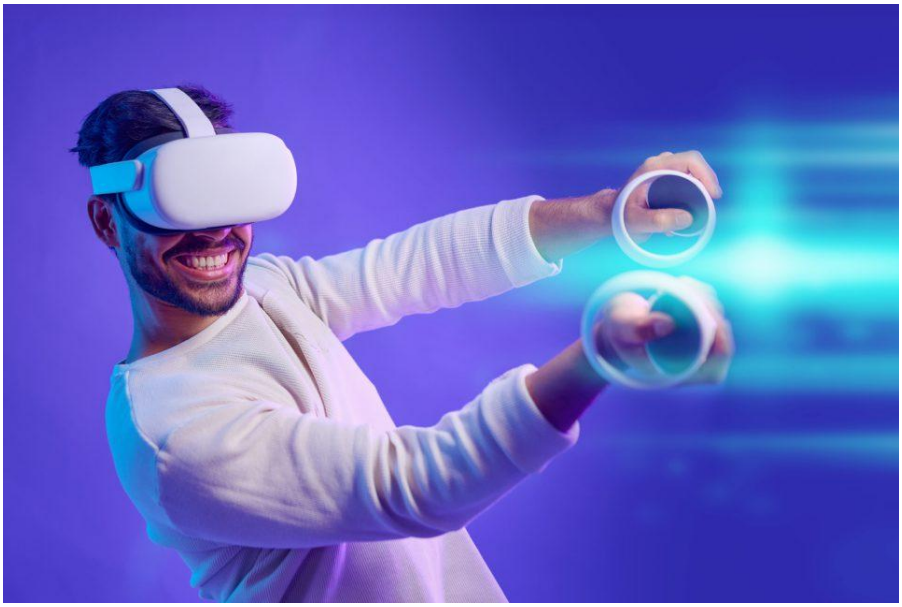


Анализата на гласот, вклучувајќи и изведување (деривација) на емоционално изразување и атрибути на говорникот, ги внесува човечките допирни точки во виртуелниот простор. Приклучокот XR им овозможува на уредите (машини) на играта да користат вокален израз како влез.

Voice AI ја подобрува интеракцијата во нови простори, создавајќи уникатно и извонредно искуство. Анализа во реално време овозможува интеграција на суштински аудитивни знаци од физичка комуникација во метаверзумот.

devAIce® XR ги нуди основните технолошки модули на audEERING како приклучок за мотори за игри (game engines):

- Димензии на емоции.
- Карактеристики на говорник.
- Откривање на гласовна активност (VAD - Voice Activity Detection).



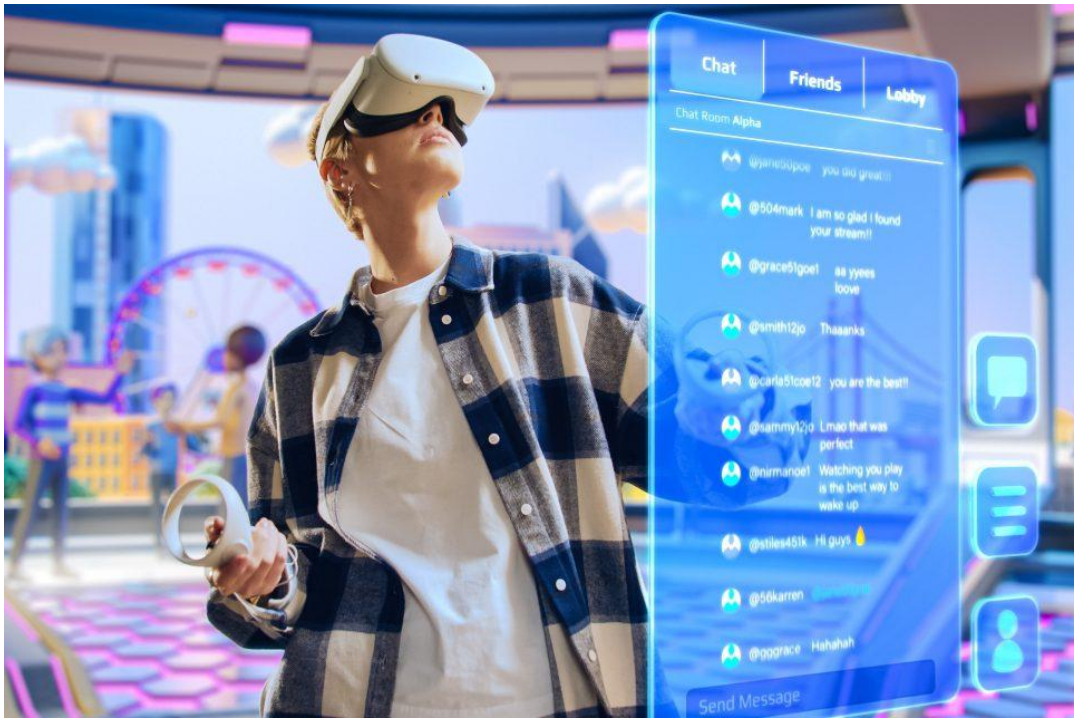
Нови креативни длабочини

Откривањето емоции овозможува неинвазивна емоционална интеракција за поголем ангажман и вклученост во XR. Користејќи вградени микрофони, devAIce XR се применува на мобилна виртуелна реалност (mobile VR) и самостојни HMD (stand-alone HMDs).

devAIce® modules integrated to kick off the XR project

Клучни карактеристики на devAIce XR.

- **Emotion Dimensions:** Valence, arousal, dominance.
- **Emotion Categories:** Happy, angry, sad, neutral (среќен, лут, тажен, неутрален).
- **Voice Activity Detection:** Снима само човечки глас.
- **Speaker Attributes:** Пол (self-reported; sex), возраст.
- **XR-Ready:** Suitable for VR, AR & MR applications.
- **Lightweight package:** Depends on activated modules. As low as 13MB.
- **Multi-platform:** Windows, Linux, MacOS, iOS, Android.
- **Memory Usage:** Depends on the activated modules. As low as 20MB.
- **Negligible CPU Usage:** Real-time factor under 1.
- **Game Engines:** Unity, Unreal.



Примена во медиција

Добра комуникација е од витално значење.. Во здравствена средина, на пример, добра комуникација помеѓу здравствените работници е важна и е познато дека има свој ефект врз севкупната безбедност на пациентите. Но, има уште повеќе. Добра комуникација помеѓу здравствен работник и пациент всушност има ефект врз самиот процес на заздравување. Истражувањата покажуваат: ако се успее да се намали анксиозноста преку начин на комуникација, пациентите ќе имаат попријатно чувство на нивниот третман. Кога анксиозноста што ја доживуваат пациентите за време на нивниот третман опаѓа, тогаш се намалува и времето што им е потребно да закрепнат. Изненадувачко? Па, очигледно не само лековите можат да бидат лековити за телото, туку и зборовите и говорот на телото.



Третман на пациенти

Човечкиот ум понекогаш го мами телото. Во медицина тоа се гледа во ефектот "**плацебо**" ('placebo' effect), што значи дека во некои случаи состојбата на пациентите се подобрува со земање лек кој не содржи никакви физиолошки активни компоненти. Трикот овде е во тоа што пациентот верува дека лекот делува, и очигледно ова верување помага во лекувањето. Неговиот пандан постои исто така, иако помалку познат, и се нарекува '**nocebo'** ефект. Истражувања покажаа дека кога пациентите очекуваат влошување на ефектот наместо лековито дејство од третманот, некои пациенти ќе доживеат влошување на нивната состојба, иако лекот што го примиле не содржел никакви активни компоненти. Очекувањата што ги имаат пациентите во однос на нивниот третман очигледно се важни за исходот, и тоа не застанува тука. Исто така, се покажа дека квалитетот на односот пациент-лекар предизвикува плацебо-одговори. Тоа ги тера некои експерти да расправаат да го заменат терминот 'placebo/nocebo' за "контекстуални ефекти" ('contextual effects').

Земањето предвид на "контекстуални ефекти" во комуникација подразбира дека треба да се избегнува содржина - зборови и говор на телото - што повикува на негативни очекувања од третманот. И наместо тоа, здравствените работници треба да ги потврдат позитивните очекувања што ги имаат пациентите. Ком-

бинирање на овој метод со сугестивен јазик, воспоставување контакт преку пресликување на говорот на телото и пренасочување на вниманието, како и други техники кои се применуваат во клиничка хипноза, може да го направат ефектот уште посилен.

Теоријата има совршена смисла, но како да се примени таа во пракса? Избегнување фрази кои можат да предизвикаат стресен одговор, како што се "болка" или "ова може малку да боли", често се користи за ментално да се подготви пациентот за претстоен третман. Иако намерата е добро замислена, ефектот е спротивен, бидејќи го тера пациентот да очекува некаква непријатност. За здравствен работник, прекршувањето на оваа "**добра намера**" и навика може да биде тешко. Зборовите треба да се избираат многу посвесно и лесно се прави лапсус. Дополнително, теоријата вели да ја пресликаме сопствената положба на телото со онаа на пациентот, бидејќи тоа би помогнало во насочувањето на разговорот на потсвесно ниво. Сепак, повеќе луѓе не се свесни за држењето на телото. Едноставно затоа што буквално не може да се гледа себеси како некој пред него. Значи, правилно прилагодување на говорот на телото на одреден пациент е тешко, па дури и може да биде контраинтуитивно.

Така, комбинирање на овие вербални и невербални техники, особено на почетокот, ќе биде прилично напорно. И ако приспособувањето на нашиот говор и држење на телото во исто време не е веќе доволно тешко, следно ниво е активно да го насочиме вербалниот и невербалниот разговор **подалеку од било каков стимул што може да имплицира стрес за пациентот**. Нешто што може да се смета за доста предизвикувачко во просечен здравствен амбиент: "Поголем број луѓе не се свесни за нивниот говор на тело".

Решение

Очигледно има потреба да се практикува ова. Во безбедна животна средина. Онаму каде што ниту еден пациент не може да биде повреден, каде што ниту еден колега не гледа зад друг. Онаму каде што се безбедни од мислења и добиваат објективни повратни информации од некој кој нема никаков интерес

освен да ги научи како да ги извршуваат и совладаат техниките. Со бескрајно трпение; може да се обидуваат повторно и повторно со свое темпо.

VR може да ни го даде токму тоа. Во проект за ко-креација во Медицинскиот центар на Универзитетот Radboud, сериозна компанија за игри, Simulation Crew се придружи со нивна експертиза за нарушена комуникација со пациенти и развој / дидактика на VR-апликации, соодветно. Резултат е решение за обука базирано на VR.

Апликацијата им овозможува на здравствените работници да практикуваат комуникациски техники за удобност на пациент во безбедно опкружување во нивна сопствена канцеларија или дом. VR нуди целосно извонредно искуство и оживува виртуелна здравствена заштита. VR очила и контролери овозможуваат да се доживее местото како навистина да се таму. Пациент за вежбање не е реа-лен, тој/таа е аватар, но реагира на она што го кажува здравствениот работник, реагира на неговото држење на телото и контакт со очите. Кога пациентот е вознемирен, неговата работа е да користи техники за **депресирање** на пациентот за да го смири. Во прво ниво, фокусот е на вербална комуникација: треба да се изберат вистински зборови за да се насочи комуникацијата. Кога ќе се направи грешка, пациентот реагира, како што би реагирал вистински пациент. По сесијата се добиваат повратни информации за тоа што е правилно и од кои точки може да се научи.

На следно ниво, треба да ги преслика позата на телото и истовремено да се користат вистински зборови. Постепено се учи да се применуваат техники на природен начин. Нивниот виртуелен пациент им дава повратни информации по секоја сесија, кажувајќи како неговиот јазик влијаел врз него, и вербално и невербално. Покрај овие повратни информации, перформанси на здравственото лице исто така се мерат на објективен начин преку збир на однапред дефинирани правила за комуникација. Ова му овозможува на вградениот тренер за виртуелно учење да им даде совети како да се подобрат. Можете да се тренира (предизвик) со други пациенти, на пример, пациент кој е лут и непријателски настроен, затоа што морал да чека долго време. Или токму спротивно: пациент кој е срамежлив и евазивен (избегнува разговор). Секој пациент по сесијата кажува како ја доживеал

интеракцијата што ја имал со лицето кое се обучува и дали неговата комуникација направила разлика. Се чини нереално? Но, не е.

Техника

RUMC Patient Comfort Techniques е пример за Iva: платформа за обука за VR комуникација.

- Природно препознавање говор,
- невербално внесување,
- откривање позиција на тело и гестови,
- виртуелни луѓе контролирани со вештачка интелигенција.

Најсовремена технологија со единствена цел да се креираат извонредни разговори, ефикасно да се обучуваат комуникациските вештини. Simulation Crew разви уникатна алатка за вежбање комуникациски вештини со виртуелни луѓе и ја нарече Iva.



“**Јас сум Ива**“

Јас сум виртуелен човек во виртуелен свет. Иако не сум реален, можеме да имаме човечки и интересен разговор. Одговарам на она што го кажуваш и ме погодува начинот на кој го кажуваш, без разлика дали остваруваш контакт со очи и говор на тело.

Заедно можеме да вежбаме безброј ситуации на разговор. Можам да преземам различни улоги, личности и емоции. Во мојот виртуелен свет сè се мери и оценува, за да добиете значајна повратна информација.

Јас сум секогаш тука, затоа вежбајте колку често и кога сакате. Ви помагам да ги подобрите вашите комуникациски вештини.

Јас сум Џојс, Макс, Ана, Пол…

Јас сум интелигентен виртуелен агент.

Јас сум Ива.

8. ПРАКТИЧНО ИСТРАЖУВАЊЕ – КОРИСТЕЊЕ СОВРЕМЕНА СЕНЗОРСКА ТЕХНОЛОГИЈА ЗА СЛЕДЕЊЕ НА ОЧИ

Истражувањето за мојата магистерска работа го реализирав во современа лабораторија за биометриски истражувања на Меѓународен Славјански Универзитет во Свети Николе, во период од 15.04.2024 до 28.04.2024 година, под менторство на проф. д-р Сашо Манасов. Во истражувањето користев напреден сензорски систем за следење на очи - хардвер и софтвер: реализирав квантитативни мерења со биосензор за следење на очи на испитаници, кои следеа визуелни стимули прикажани на компјутерски екран, за да се процени визуелното внимание (на кои елементи од стимулот испитаниците покажуваат најголем интерес.

Следењето на очите на испитаници е современа технологија која овозможува мерење на визуелното внимание, метрика која следи **ШТО**, **КАДЕ** и **КОГА** гледаат избрани испитаници. Во оваа истражување ги реализирав следните метрики преку анализа на добиените мерења:

- **Топлински мапи** (**heat maps**): прикажуваат дистрибуција на визуелно внимание кое е пропорционално на нивото на интерес.
- **Области на интерес** (**Areas of interest – AOI**): Однапред дефинирани и креирани бафер зони за одредени области од екранскиот стимул.

- **Низи на фиксации (fixation sequences)**: локација од стимулот каде гледаат испитаниците и колку време го задржале погледот во гледање определен елемент од стимулот.
- **Дилатација на зеница.** Проширувањето на зениците на очите е пропорционална на **ЈАЧИНАТА НА ЕМОЦИОНАЛНА ВОЗБУДА**.
- **Фреквенција на трепкање:** Ниска фреквенција (помал број на трепнувања од испитаник за време на траење на опсервација на екрански стимул – 10 секунди) е поврзана со високо ниво на концентрација, а висока фреквенција на трепкање е индикативна за поспаност, пониско ниво на фокус и пониско ниво на концентрација.

Во ова истражување го користев следниот стимул:

- **Стимул**: Избор на риба – конзерва со дефинирана цена на чинење.

Преку ова истражување покажувам каков избор прават испитаниците од понудени 4 различни производи, и како цената на чинење (иста цена за сите производи) влијае на нивниот избор. Притоа еден производ е потенциран, предложен (цената е со жолто подвлечена), за да се види дали испитаникот ќе го избере производот што е предложен од понудувачот.

Реализирав два (2) различни вида на истражување:

- Биосензорско следење на очи со тракер базиран на инфрацрвена светлина.
- Анкетно истражување – заокружување (избор) кој производ најмногу им се допаднал.

Анализа на резултатите од Биосензорско следење на очи

a) **Топлински мапи**: **Поединечни** анализи

Црвени области: Покажуваат зголемено ниво на интерес.

Жолти и зелени области: Покажуваат помало ниво на интерес.

Области кои не се обоени, не се од интерес на испитаниците.

Во истражувањето е користен следниот екрански стимул:



Слика. Стимул (слика) за опсервација со екрански биосензор за следење на очи и анкетно истражување

Цел на биосензорско мерење: Да се измери визуелното внимание (каде најмногу го задржале погледот) на група испитаници кога следат одреден стимул – избор на риба-конзерва за купување, од понудени четири различни видови.

8.1 Анализа на резултатите од биосензорско следење на очи: поединечни и збирни анализи

Пред да се започне со следење на очите на секој испитаник, се реализира процес на **Калибрација** (за секој испитаник посебно) – прилагодување за да се добијат најточни и целосни резултати.

На слика 1 е прикажана топлинска мапа на испитаник 3 при следење на екрански стимул. Се забележува дека овој испитаник најголем интерес покажа за првиот производ, додека нешто помал интерес за другите производи. На анкетниот прашалник, овој испитаник го заокружил (избрал) производ 1, што покажува дека не постои дискрепанца (отстапување) помеѓу двата методи на истражување (Табела на крајот на трудот). Биосензорското следење на очи го покажува потсвесното размислување на испитаникот, додека анкетното истражување неговото рационално размислување.



Слика 1. Топлинска мапа на испитаник 3

На слика 2 е прикажана топлинска мапа (spotlight style), каде интензитетот на осветленост е индикатор, пропорционален на фокусирано внимание. Осветленоста е најголема кај производ 1 (најголем интерес), и се забележува дека овој испитаник покажал интерес и за цената на чинење (една од тие е и онаа што е подвлечена жолто).

На слика 3 е прикажана топлинска мапа (shadow style), каде површината на сенката е индикатор, пропорционален на фокусирано внимание. Површината на сенката е најголема кај производ 1 (најголем интерес), и е нешто помала кај другите производи.



Слика 2. Топлинска мапа (spotlight), испитаник 3



Слика 3. Топлинска мапа (shadow), испитаник 3

На слика 4 и слика 5 се прикажани топлински мапи за испитаник 3, каде дополнително се прикажани поединечни фиксации (погледи) (грозд–плави точки).



Слика 4. Точки на визуелизација (spotlight), испитаник 3



Слика 5. Точки на визуелизација (spotlight), испитаник 3

На слика 6 е прикажана топлинска мапа на испитаник 9. Се забележува дека биосензорското мерење покажува дека овој испитаник најголем интерес покажал за производот 3. За другите производи тој интерес е помал. Но, на анкетниот прашалник испитаникот го заокружил (одбрал) производ 1. Ова покажува дека постои дискрепанца (отстапување) помеѓу двата методи на истражување. Ова отстапување ги покажува следните два вида размислување кај овој испитаник:

- Со биосензорско мерење се регистрира **потсвесното** размислување на испитаникот. Тоа е брзо, повеќе емоционално размислување. Ако на овој испитаник се даде кратко време за размислување, тогаш тој ќе го избере производ 1.
- Со анкетниот прашалник се регистрира **рационалното** размислување на испитаникот. Тоа е бавно размислување, логичко и ослободено од емоции. Ако на овој испитаник се даде подолго време за размислување, тогаш тој ќе го избере производ 3.



Слика 6. Топлинска мапа, испитаник 9



Слика 7. Топлинска мапа (shadow), испитаник 9

На слика 7 е прикажана топлинска мапа (shadow style), каде површината на сенката е индикатор, пропорционален на фокусирано внимание. Површината на сенката е најголема кај производ 3 (најголем интерес), и е нешто помала кај другите производи.

На слика 8 е прикажана топлинска мапа (spotlight style), каде интензитетот на осветленост е индикатор, пропорционален на фокусирано внимание. Осветленоста е најголема кај производ 3 (најголем интерес), и се забележува дека овој испитаник покажал интерес и за цената на чинење (една од тие е и онаа што е подвлечена жолто). Исто така испитаникот покажал интерес и за другите производи. Интересно е созднанието дека испитаникот ги набљудувал и цените на чинење на трите производи, освен на третиот производ (за кој најголем дел од времето на набљудување) за кој покажал најголем интерес



Слика 8. Топлинска мапа (spotlight), испитаник 9



Слика 9. Топлинска мапа (spotlight), испитаник 9, со прикажани групирани погледи на очите

На слика 9 и слика 10 се прикажани топлински мапи за испитаник 9, каде дополнително се прикажани поединечни фиксации (погледи) (грозд–плави точки). Се забележува дека на производот 3 има најголемо групирање на поединечни фиксации на погледот (конвергенција), додека кај другите има нивно расејување (дивергенција).



Слика 10: Топлинска мапа на испитаник 9, со групиран поглед

b) Топлински мапи (Heat maps): збирни анализи

Збирни топлински мапи се најдобар показател за преференциите на одредена таргет група (испитаници), бидејќи ги презентира нивните збирни одлуки: кој производ е најпосакуван за нив. Оваа метрика најчесто се користи во маркетинг истражувања, каде на носителите на маркетинг стратегии (маркетери) им дава егзактни податоци во процесот на донесување одлуки. Биосензорските збирни анализи денес ги користат многу компании во светот: имаат свои невромаркетинг лаборатории или користат услуги на истражувачки лаборатории за да добијат

вакви релевантни и многу корисни информации (како размислуваат потрошувачите – одредени целни групи).



Слика 11. Топлинска мапа, збирна (од сите испитаници)

На слика 11 е прикажана збирна топлинска мапа. Може да се забележи дека испитаниците најголем интерес покажале за производ 3 (пет од десет испитаници го преферирале овој производ преку биосензорско следење на очите – 50% од испитаниците). За производот 1 се одлучиле 4 од 10 испитаници (40% од испитаниците). Ова е прикажано во Табела 1 на крајот на трудот.

На анкетниот прашалник најголем број испитаници (5 од 10) го избрале (заокружиле) производ 1. Тоа е 50% од вкупниот број испитаници. За производот 3 се изјасниле 30% од испитаниците (3 од 10 испитаници).

Овие резултати покажуваат за потсвесното и рационалното размислување на испитаниците. Ова е значаен податок за менаџерите за маркетинг стратегии, особено во процесот на рекламирање: кога да се користат кратки, визуелни и брзи реклами (потсвесно размислување), а во кои услови да се користат пишани реклами (или online пораки) за рационално, бавно размислување на потрошувачите.



Слика 12. Топлинска мапа, збирна (од сите испитаници)

На слика 12 е прикажана збирна топлинска мапа (spotlight style), каде интензитетот на осветленост е индикатор, пропорционален на фокусирано внима-ние. Осветленоста е најголема кај производ 3 (најголем интерес). Исто така испитаниците покажале интерес и за другите производи. Интересно е созднанието дека испитаниците ги набљудувале и цените на чинење на трите производи, мајмногу на четвртиот производ која е потенцирана со жолто.

На слика 13 е прикажана збирна топлинска мапа (shadow style), каде површината на сенката е индикатор, пропорционален на фокусирано внимание. Површината на сенката е најголема кај производ 3 (најголем интерес), и е нешто помала кај другите производи.

На слика 14 и слика 15 се прикажани збирни топлински мапи, каде дополнително се прикажани поединечни фиксации (погледи) (грозд–плави точки). Се забележува дека на производот 3 има најголемо групирање на поединечни фиксации на погледот (конвергенција), а кај другите има нивно расејување (дивергенција).



Слика 13. Топлинска мапа, збирна (од сите испитаници)



Слика 14. Топлинска мапа, збирна (од сите испитаници), со групирани погледи на очите



Слика 15. Топлинска мапа, збирна (од сите испитаници)

c) Низи на фиксации (Fixation sequences)

Низи на фиксации се генерираат врз основа на:

- **Позицијата на фиксацијата**:**КАДЕ е насочен погледот**.
- **Информации за времето: КОГА** е реализиран погледот.

Покажува каде испитаниците погледнале: редослед на гледање.

Ова е една многу често користена метрика во истражување со биосензор за следење на очи.

Низа од фиксации се реализира **за секој испитаник посебно**, при опсервација на посебен **стимул**.

На слика 16 се прикажани низа на фиксации на испитаник 3. Се забележува дека дилатацијата на зеницата на левото и десното око, е во одредени граници. Вредностите на дилатацијата на зеницата раскажуваат како испитаникот се чувствува (ниво на возбуда) кога гледа одредени сегменти од стимулот. Дилатацијата на зеницата е пропорционална на нивото на возбуда на испитаникот.

Прикажана е и фреквенција на трепкање на испитаникот (едно трепкање – син столб на сликата). Оваа ниска фреквенција на трепкање индицира високо ниво на концентрација на испитаникот кога го набљудувал екранскиот стимул.

Вредностите на зениците континуирано се прикажуваат за време на целиот период на опсервација на стимулот (10 секунди).



Слика 16: Низа на фиксации на испитаник **3**

На слика 17 е прикажана низа од фиксации на испитаник 3 само во првата секунда кога го набљудувал стимулот (како го започнал своето набљудување). Со анализа на резултатите од биосензорско следење може да се прави анализа каде гледал испитаникот во произволен период од неговото следење и да се следи нивото на возбуда.

Од оваа визуелна анализа може да се заклучи дека, испитаникот 3 во првата секунда од следењето на стимулот, своето внимание го насочува на првите три производи, т.е неговото набљудување се движи од лево кон десно.

Исто така може да се заклучи дека неговиот интерес на почетокот не е поврзан со потенцираната цена на чинење (подвлечено со жолто), иако тоа било

целта на трговецот (да потенцира одредена цена со што да побуди зголемен интерес на потрошувачите да го купат тој производ).



Слика 17: Низа на фиксации на испитаник **3** (фиксации во првата секунда од следењето на стимулот)

На следната визуелна анализа (слика 18) се забележува дека потенцираната цена на чинење побудила интерес кај испитаникот кога поминала 1.5 секунди од почетокот на следењето на стимулот.

На слика 19 е прикажана низа од фиксации на испитаник 9 во првата секунда од следење на стимулот. Се забележува каде бил неговиот интерес на почеток од набљудувањето.

Од оваа визуелна анализа може да се заклучи дека, испитаникот 9 во првата секунда од следењето на стимулот, своето внимание го насочува на првите два производи, т.е неговото набљудување се движи од лево кон десно. Исто така може да се заклучи дека неговиот интерес на почетокот не е поврзан со потенцираната цена на чинење (подвлечено со жолто)



Слика 18: Низа на фиксации на испитаник **3** (фиксации во првите 1.5 секунди од следењето на стимулот)



Слика 19: Низа на фиксации на испитаник **9** (фиксации во првата секунда од следењето на стимулот)

На следната визуелна анализа (слика 20) се забележува дека потенцираната цена на чинење побудила интерес кај испитаникот кога поминала 2.4 секунди од почетокот на следењето на стимулот.

Во времето до 4 секунди (слика 21) од почетокот на следењето на стимулот, испитаникот значајно време посветува на четвртиот производ (зголемено ниво на интерес), што значи дека потенцираната цена на чинење го привлекла неговото внимание.



Слика 20: Низа на фиксации на испитаник **9** (фиксации во првите 2.4 секунди од следењето на стимулот)

Исто така, апликативното софтверско решение овозможува реализирање на анализи со кои се генерира визуелно претставување на фиксации (за било кој испитаник и за било кој стимул) во произволен временски интервал. Ова овозможува да се набљудува однесувањето на испитаниците во специфични временски интервали, при карактеристично ниво на проширување на зеници и фреквенција на трепкање (одредени емоционални моменти: концентрација, фокусираност,

поспаност, ниво на интерес, јачина на емоционална возбуда, когнитивно оптоварување...).



Слика 21: Низа на фиксации на испитаник **9** (фиксации во првите 4.0 секунди од следењето на стимулот)

Низа на фиксации може да се прикажуваат (анализираат) на повеќе начини:

- **Динамички**: движење на фиксации како што се генерирале (редослед на појавување – од прва до последна).
- **Статички**: сите фиксации одеднаш (во исто време) се појавуваат на стимулот).
- **Рачно:** рачно се влече временска линија со што се овозможува детално анализирање на набљудувачкиот процес - се гледа која фиксација кога се појавила, колку траела и каде се генерирала.

Проширување на зеница (pupil dilation): Претставува метрика која е интегрирана во биосензорскиот уред - Eye tracker, и не се појавува во основните табели, туку се генерира во посебни табели (истото важи и за метриката Трепкања -

Blinks). Од промената на големината на зеницата (процес на дилатација), може да се изведат и други важни атрибути, значајни за човековото однесување:

- **Емоционална возбуда** (emotional arousal): Количество на емоционална ангажираност.
- **Когнитивен обем на работа** (cognitive workload): Какво ментално оданочување е стимулот.

Можат да се прикажат следните форми (дилатација):

- Лева зеница (Left pupil).
- Десна зеница (Right pupil).
- Комбинирана големина на двете зеници (Combined pupil size).

Проширување на зеница се прикажува за секој испитаник оделно, и за секој стимул оделно. Проширувањето на зеницата всушност го мери дијаметарот на зеницата на левото и десното око посебно.

d) **Области на интерес** (Areas of interest): **Поединечни** анализи

Области на интерес претставуваат кориснички дефинирани **подрегиони** на прикажан стимул. Генерирањето метрики за одделни **AOI** овозможува да се прават **проценки на перформанси** (evaluating the performance) на две или повеќе специфични области на иста слика (стимул). На ваков начин се реализираат **СПОРЕДБИ** (пример, на два или повеќе производи) и други **статистички** и **математички** анализи на иста сцена.

При креирање на одредена област од интерест, се покажуваат повеќе метрики кои даваат карактеристики за визуелното внимание за одреден стимул и за секој поединечен испитаник.

На слика 22 е прикажана област на интерес за испитаник 3, при што се генерирани четири (4) бафер зони околу секој производ, за да се добијат повеќе метрики кои ќе овозможат да се анализира нивото на интерес на испитаникот, нивото на фокусирано внимание и неговото однесување (емоционална возбуда).

Овој испитаник најголем интерес покажал за производ 1 (25.5% од времето на набљудување го поминал следејќи го овој производ). Нешто помал интерес покажал за производ 2 – риба 2 (21.5% поминато време на следење), додека значајно помал интерес покажал за останатите два производи.

На анкетниот прашалник овој испитаник го заокружил (избрал) првиот производ (риба 1) што покажува дека не постои дискрепанца (отстапување) во двата методи на истражување. Ова значи дека при било каков процес на размислување (брзо-потсвесно или бавно-рационално) овој испитаник ќе го избере првиот производ.



Слика 22. Област на интерес, испитаник 3

На слика 23 е прикажана иста област на интерес (како и слика 22) на испитаник 3, само што за секој производ се прикажани топлински мапи и збир на поединечни погледи (грозд на точки – плави точки) кои покажуваат на каков начин испитаникот го фокусирал своето внимание при следење на екранскиот стимул.



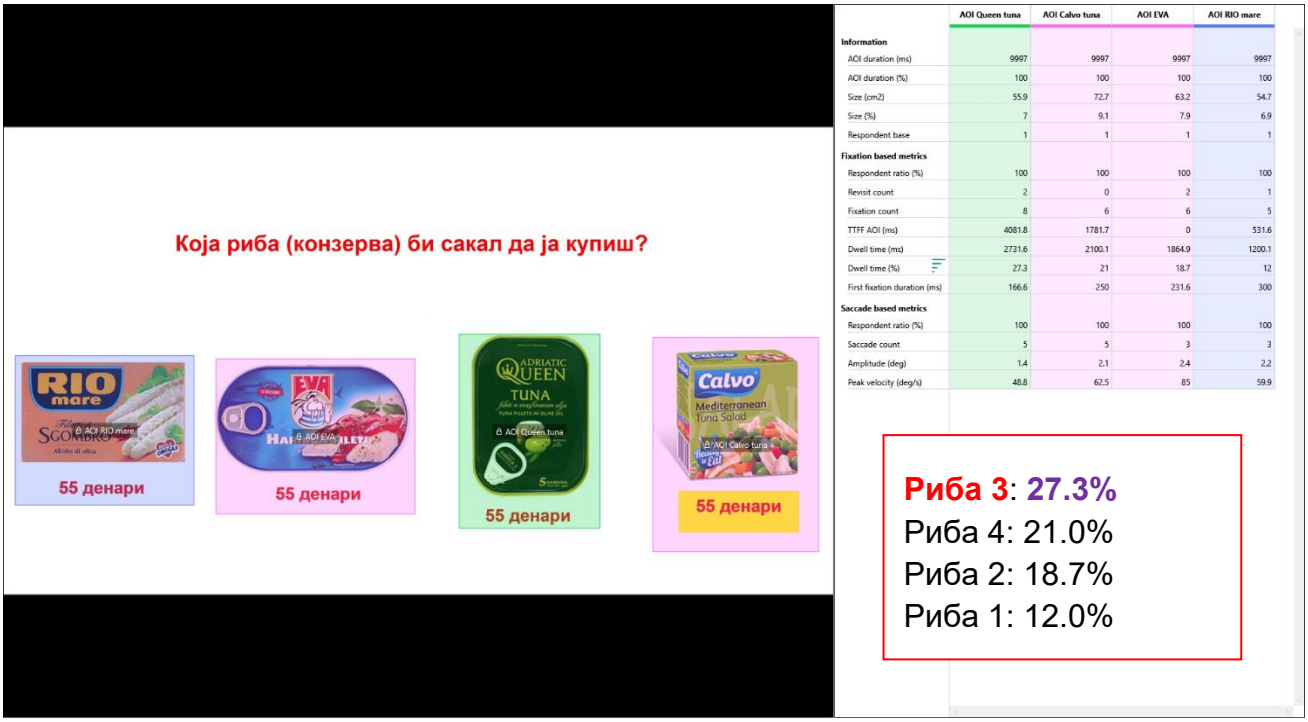
	AOI RIO mare	AOI EVA	AOI Calvo tuna	AOI Queen tuna
Information				
AOI duration (ms)	10000	10000	10000	10000
AOI duration (%)	100	100	100	100
Size (cm2)	54.7	63.2	72.7	55.9
Size (%)	6.9	7.9	9.1	7
Respondent base	1	1	1	1
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	3	5	3	6
Fixation count	11	9	6	8
TTFF AOI (ms)	0	501.3	1368	1001.3
Dwell time (ms)	2551.6	2150.1	1516.8	1381.6
Dwell time (%)	25.5	21.5	15.2	13.8
First fixation duration (ms)	368	216.7	200	133.4
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	8	6	3	2
Amplitude (deg)	2.3	2.6	2.8	3.9
Peak velocity (deg/s)	79.4	61.4	100.1	110.9

Слика 23. Област на интерес, испитаник 3 (со топлинска мапа и точки на поглед)

На слика 24 е прикажана област на интерес на испитаник 9. Овој испитаник најголем интерес покажал за **производ 3** (**27.3%** од времето на набљудување го поминал следејќи го овој производ). Нешто помал интерес покажал за производ 4 – риба 4 (21.0% поминато време на следење), додека значајно помал интерес покажал за останатите два производи.

На анкетниот прашалник овој испитаник го заокружил (избрал) првиот производ (риба 1) што покажува дека постои дискрепанца (отстапување) во двата методи на истражување. Ова значи дека:

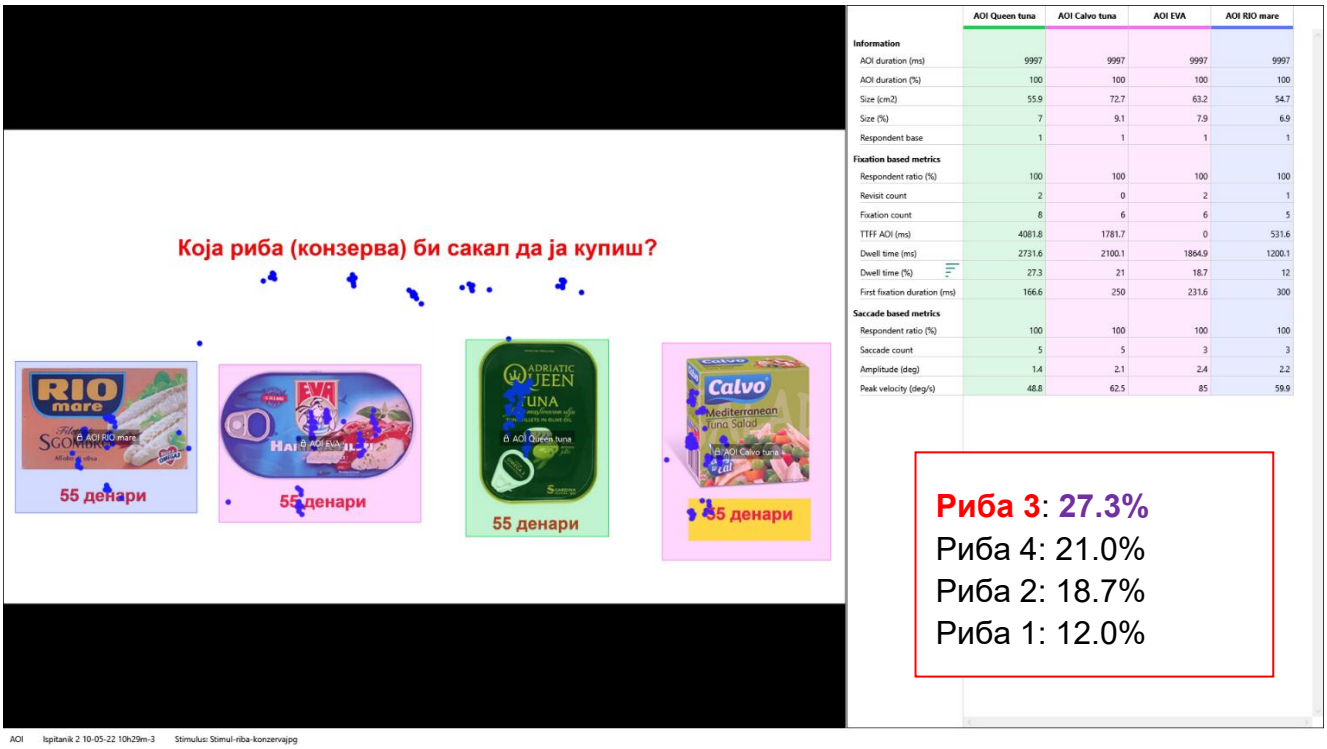
- При процес на брзо донесување одлука (потсвесно размислување, регистрирано со биосензор за следење очи), испитаникот 9 ќе го избере производ 3.
- При процес на бавно донесување одлука (рационално размислување, регистрирано со анкетно истражување), испитаникот 9 ќе го избере производ 1.



	AOI Queen tuna	AOI Calvo tuna	AOI EVA	AOI RIO mare
Information				
AOI duration (ms)	9997	9997	9997	9997
AOI duration (%)	100	100	100	100
Size (cm2)	55.9	72.7	63.2	54.7
Size (%)	7	9.1	7.9	6.9
Respondent base	1	1	1	1
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	2	0	2	1
Fixation count	8	6	6	5
TTFF AOI (ms)	4081.8	1781.7	0	531.6
Dwell time (ms)	2731.6	2100.1	1864.9	1200.1
Dwell time (%)	27.3	21	18.7	12
First fixation duration (ms)	166.6	250	231.6	300
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	5	5	3	3
Amplitude (deg)	1.4	2.1	2.4	2.2
Peak velocity (deg/s)	48.8	62.5	85	59.9

AOI Ispitanik 2 10-05-22 10h29m-3 Stimulus: Stimul-riba-konzervajpg

Слика 24. Област на интерес, испитаник 9

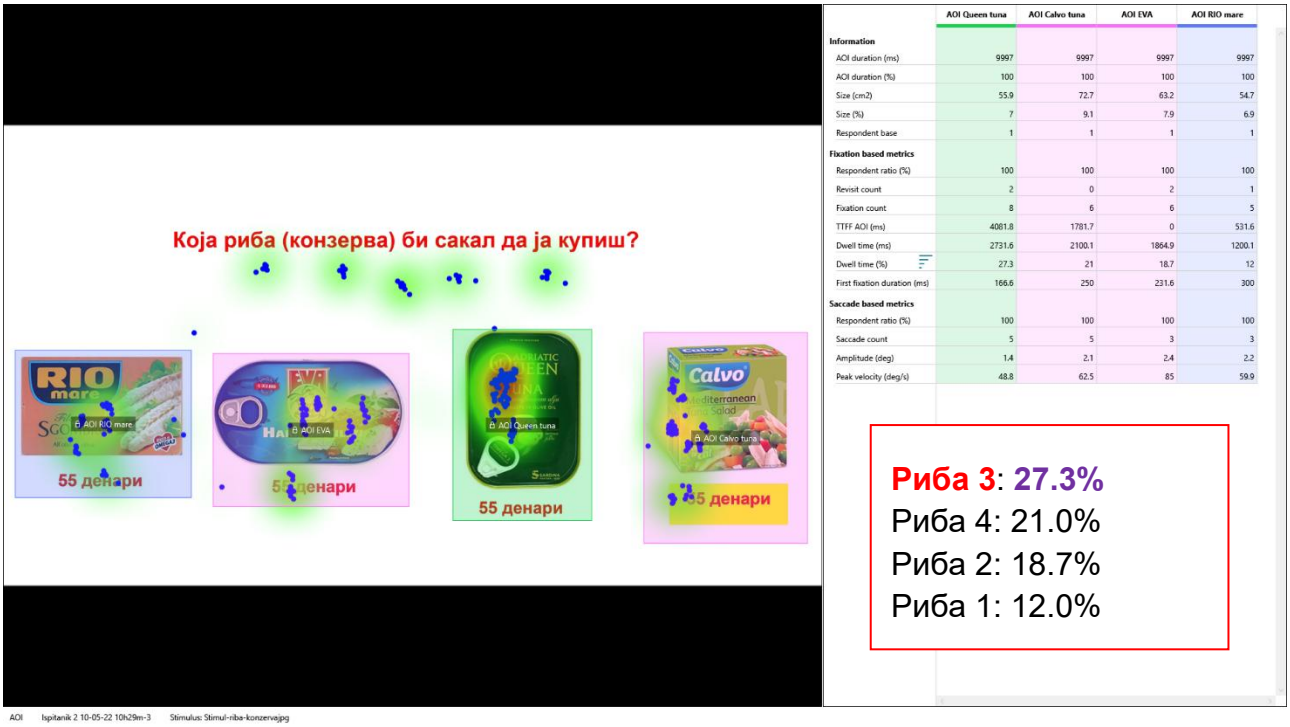


	AOI Queen tuna	AOI Calvo tuna	AOI EVA	AOI RIO mare
Information				
AOI duration (ms)	9997	9997	9997	9997
AOI duration (%)	100	100	100	100
Size (cm2)	55.9	72.7	63.2	54.7
Size (%)	7	9.1	7.9	6.9
Respondent base	1	1	1	1
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	2	0	2	1
Fixation count	8	6	6	5
TTFF AOI (ms)	4081.8	1781.7	0	531.6
Dwell time (ms)	2731.6	2100.1	1864.9	1200.1
Dwell time (%)	27.3	21	18.7	12
First fixation duration (ms)	166.6	250	231.6	300
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	5	5	3	3
Amplitude (deg)	1.4	2.1	2.4	2.2
Peak velocity (deg/s)	48.8	62.5	85	59.9

AOI Ispitanik 2 10-05-22 10h29m-3 Stimulus: Stimul-riba-konzervajpg

Слика 25. Област на интерес, испитаник 9 (со точки на поглед)

На слика 25 и слика 26 се прикани област на интерес за испитаник 9 како на слика 24, само што кај овие бафер зони на производите се прикажани топлински мапи и точки на поглед.



	AOI Queen tuna	AOI Calvo tuna	AOI EVA	AOI RIO mare
Information				
AOI duration (ms)	9997	9997	9997	9997
AOI duration (%)	100	100	100	100
Size (cm2)	55.9	72.7	63.2	54.7
Size (%)	7	9.1	7.9	6.9
Respondent base	1	1	1	1
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	2	0	2	1
Fixation count	8	6	6	5
TTFF AOI (ms)	4081.8	1781.7	0	531.6
Dwell time (ms)	2731.6	2100.1	1864.9	1200.1
Dwell time (%)	27.3	21	18.7	12
First fixation duration (ms)	166.6	250	231.6	300
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	5	5	3	3
Amplitude (deg)	1.4	2.1	2.4	2.2
Peak velocity (deg/s)	48.8	62.5	85	59.9

Слика 26. Област на интерес, испитаник 9 (со точки на поглед и топлинска мапа)

е) Области на интерес (Areas of interest): **збирни** анализи

На слика 27 е прикажана збирна топлинска мапа со прикажана област на интерес генерирана со четири (4) бафер области околу секој производ одделно. Може да се забележи каков интерес покажале испитаниците за производите кои се понудени. Најголем интерес покажала **риба 3** (**20.6%**), понатаму риба 2 (20.1%), риба 1 (18.8%) и риба 4 (18.1%). Ова покажува дека ниту еден производ не се издвојува како најдобар, најпосакуван од испитаниците, туку сите четири се со блиски метрики за фокусирано внимание. Ако се погледаат збирните квантитативни вредности (Табела 1) при избор на производ од сите испитаници, се констатира следново:

- **Биосензорско следење на очи**: **Риба 3** (**50.0%**) и риба 1 (40.0%).

- **Анкетно истражување**: **Риба 1** (**50.0%**) и риба 3 (30.0%).



	AOI Queen tuna	AOI EVA	AOI RIO mare	AOI Calvo tuna
Information				
AOI duration (ms)	9998.5	9998.5	9998.5	9998.5
AOI duration (%)	100	100	100	100
Size (cm2)	55.9	63.2	54.7	72.7
Size (%)	7	7.9	6.9	9.1
Respondent base	2	2	2	2
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	4	3.5	2	1.5
Fixation count	8	7.5	8	6
TTFF AOI (ms)	2541.5	250.6	265.8	1574.9
Dwell time (ms)	2056.6	2007.5	1875.8	1808.5
Dwell time (%)	20.6	20.1	18.8	18.1
First fixation duration (ms)	150	224.2	334	225
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	3.5	4.5	5.5	4
Amplitude (deg)	2.7	2.5	2.3	2.5
Peak velocity (deg/s)	79.8	73.2	69.7	81.3

Слика 27. Област на интерес, збирна анализа со топлинска мапа (од сите испитаници)

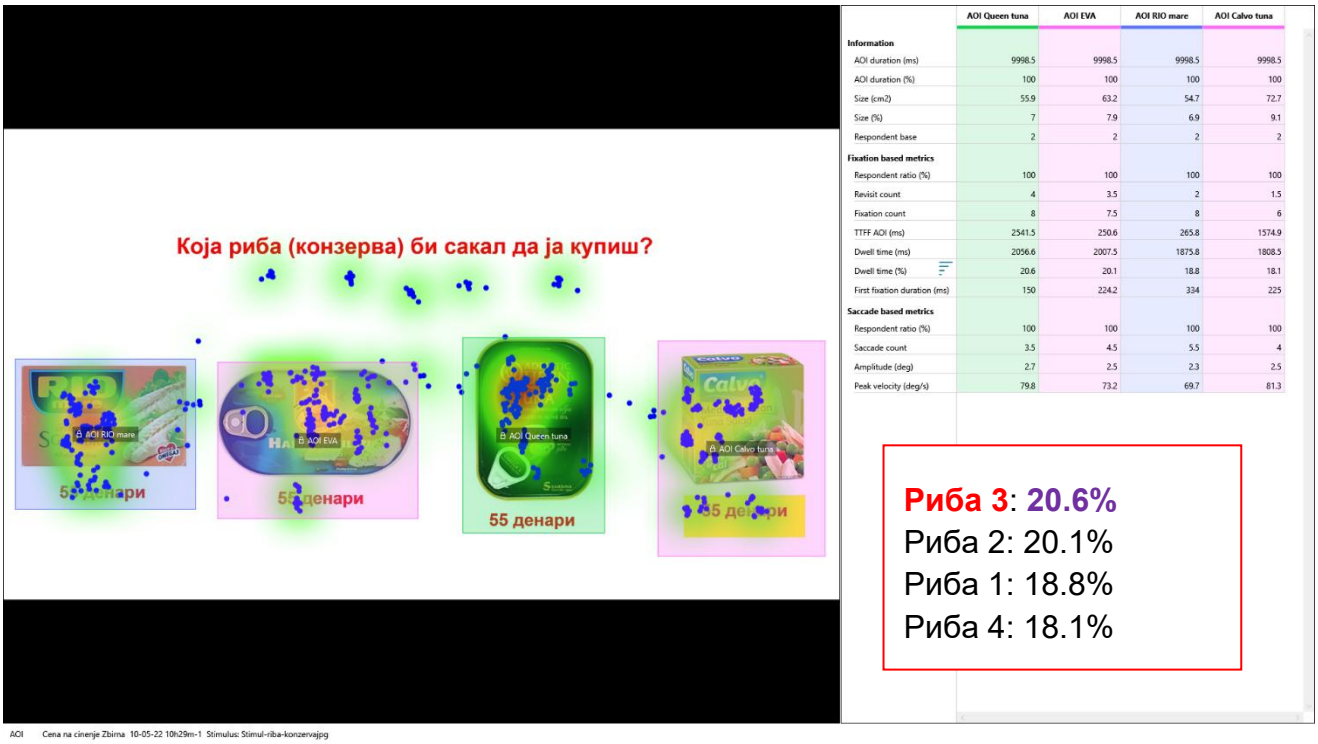
На слика 28 е прикажана иста топлинска мапа како и слика 27, само овде се прикажани кластери (грозд) на поединечни фиксации (сини точки) за секој производ поединечно. Распоредот на поединечни фиксации покажува на кои сегменти од секој производ испитаниците го насочиле својот поглед, а со тоа го покажале и каков е нивниот интерес.

На слика 29 е прикажана слична топлинска збирна мапа како и слика 28 со дополнителен приказ на стандардна топлинска мапа (бои кои го покажуваат нивото на интерес од сите испитаници заедно. Исто така се прикажани **кластери** (**грозд**) на поединечни фиксации (сини точки) за секој производ поединечно. Овој визуелен приказ е одличен начин да се покаже корелација помеѓу фокусирано внимание прикажано со бои и поединечни фиксации кои формираат кластер од збирни поединечни фиксации од секој испитаник.



	AOI Queen tuna	AOI EVA	AOI RIO mare	AOI Calvo tuna
Information				
AOI duration (ms)	9998.5	9998.5	9998.5	9998.5
AOI duration (%)	100	100	100	100
Size (cm2)	55.9	63.2	54.7	72.7
Size (%)	7	7.9	6.9	9.1
Respondent base	2	2	2	2
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	4	3.5	2	1.5
Fixation count	8	7.5	8	6
TTFF AOI (ms)	2541.5	250.6	265.8	1574.9
Dwell time (ms)	2056.6	2007.5	1875.8	1808.5
Dwell time (%)	20.6	20.1	18.8	18.1
First fixation duration (ms)	150	224.2	334	225
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	3.5	4.5	5.5	4
Amplitude (deg)	2.7	2.5	2.3	2.5
Peak velocity (deg/s)	79.8	73.2	69.7	81.3

Слика 28. Област на интерес, збирна анализа со точки на поглед (од сите испитаници)



	AOI Queen tuna	AOI EVA	AOI RIO mare	AOI Calvo tuna
Information				
AOI duration (ms)	9998.5	9998.5	9998.5	9998.5
AOI duration (%)	100	100	100	100
Size (cm2)	55.9	63.2	54.7	72.7
Size (%)	7	7.9	6.9	9.1
Respondent base	2	2	2	2
Fixation based metrics				
Respondent ratio (%)	100	100	100	100
Revisit count	4	3.5	2	1.5
Fixation count	8	7.5	8	6
TTFF AOI (ms)	2541.5	250.6	265.8	1574.9
Dwell time (ms)	2056.6	2007.5	1875.8	1808.5
Dwell time (%)	20.6	20.1	18.8	18.1
First fixation duration (ms)	150	224.2	334	225
Saccade based metrics				
Respondent ratio (%)	100	100	100	100
Saccade count	3.5	4.5	5.5	4
Amplitude (deg)	2.7	2.5	2.3	2.5
Peak velocity (deg/s)	79.8	73.2	69.7	81.3

Слика 29. Област на интерес, збирна анализа со точки на поглед и топлинска мапа (од сите испитаници)

1. **Табела 1: Споредување на поединечни вредности за сите испитаници**

Стимул			
Испитаник	**Анкетен прашалник**	**Биометриско мерење со Eye Tracker**	**Дискрепанца (2-3)**
Испитаник **1**	**Риба-конзерва 1**	**Риба-конзерва 1**	**нема**
Испитаник **2**	**Риба-конзерва 3**	**Риба-конзерва 3**	**нема**
Испитаник **3**	**Риба-конзерва 1**	**Риба-конзерва 1**	**нема**
Испитаник **4**	**Риба-конзерва 1**	**Риба-конзерва 3**	**Има: 1/3**
Испитаник **5**	**Риба-конзерва 3**	**Риба-конзерва 3**	**нема**
Испитаник **6**	**Риба-конзерва 4**	**Риба-конзерва 1**	**Има: 4/1**
Испитаник **7**	**Риба-конзерва 2**	**Риба-конзерва 2**	**нема**
Испитаник **8**	**Риба-конзерва 1**	**Риба-конзерва 1**	**нема**
Испитаник **9**	**Риба-конзерва 1**	**Риба-конзерва 3**	**Има: 1/3**
Испитаник **10**	**Риба-конзерва 3**	**Риба-конзерва 3**	**нема**
Збирни резултати	**Риба-кон. 1: 5/10 (50.0%)** **Риба-кон.2: 1/10 (10.0%)** **Риба-кон.3: 3/10 (30.0%)** **Риба-кон.4: 1/10 (10.0%)**	**Риба-кон.1: (40.0.0%)** **Риба-кон.2: (10.0%)** **Риба-кон.3: (50.0%)** **Риба-кон.4: (0.0%)**	

ЗАКЛУЧНИ СОГЛЕДУВАЊА

Едноставно кажано, следењето на очите се однесува на мерење на активноста на очите. Поконкретно, следењето на очите го опишува снимањето на положбата и движењето на очите во средина заснована на оптичко следење на рефлексиите на рожницата за да се процени визуелното внимание. Иако идејата за следење очи е прилично јасна, технологијата зад неа е прилично сложена и недостапна.

Примена на технологија за следење очи е во пораст. Додека првите уреди беа многу наметливи и вклучуваа особено незгодни процедури за поставување, современите тракери за очи претрпеа значајна технолошка еволуција во последниве години. Одамна ги нема цврстите експериментални поставки и распоредот на седишта.

Модерните тракери за очи се едвај поголеми од паметните телефони и обезбедуваат исклучително природно искуство за испитаниците.

Далечинските, ненаметливи методи го направија следењето на очите и лесно за употреба и достапна алатка во истражувањето на човековото однесување што овозможува објективни мерења на движењата на очите во реално време.

Повеќето модерни тракери за очи користат блиска инфрацрвена технологија заедно со камера со висока резолуција (или друг оптички сензор) за следење на насоката на погледот. Основниот концепт, вообичаено се нарекува **Рефлексија на рожницата на зеницата**.

Тоа во суштина вклучува: камера што го следи центарот на зеницата и каде светлината се рефлектира од рожницата. Сликата за тоа како изгледа ова е она витинското. Математиката зад неа е малку посложена. Но, овде нема да се разгледува природата на алгоритмите.

Традиционалните методи кои долго време се користат за истражување на човековото однесување се фокусираат на пристап на одговор на прашање (било да е тоа со анкети, прашалници или преку едноставно набљудување). Иако овој пристап може да генерира голем увид, на моменти може да недостасуваат кванти-

тативни и објективни податоци потребни за поддршка или побивање на изнесените тврдења.

Постојат различни причини зошто пристапот на прашање-одговор има свои недостатоци. Таквите методи се склони кон **пристрасност**, со испитаници кои не сакаат или не можат да го кажат она што го мисли нивниот ум. Дури и кога испитаниците се обидуваат да бидат целосно отворени и искрени, во игра се **потсвесни пристрасности** (subconscious biases) и разни други фактори кои можат да го променат нивниот одговор. Со обезбедување метод за мерење на потсвесните мисли и чувства на испитаник, овие фактори можат да се квантифицираат, споредуваат и искористат за да се направат заклучоци што можат да ги поддржат или побијат оригиналните хипотези. Секој биосензор обезбедува различен и колаборативен елемент во тоа разбирање.

Традиционалните методи и биосензорните методи можат да се надополнуваат одговарајќи на различни типови прашања во врска со психолошкото и физиолошкото патување на една личност во однос на одреден стимул или искуство.

Голем дел од човековото однесување е водено од несвесни процеси. Деценициски истражувања покажаа дека голем дел од нашата ментална обработка се случува на потсвесно ниво. Ова може да се подобри со користење апликативен интеграбилен софтвер кој им дава на истражувачите пристап до **емоционални**, **когнитивни** и **бихејвиорални (**однесување**)** податоци. Ваков софтвер интегрира и синхронизира повеќе **БИОМЕТРИСКИ СЕНЗОРИ** кои овозможуваат различен увид на човечки **психо-физиолошки** појави.

Говорниот јазик е нашиот избран тип на комуникација уште од нашите рани денови како вид. Всушност, веројатно е многу подлабоко во нашата еволутивна историја од првите луѓе кои се појавија. Звуците се толку длабоко вкоренети во начинот на кој комуницираме со нашиот свет што е невозможно да се размислува за иднина во која нашата комуникација со машини не е барем делумно заснована на овој модалитет.

Крвта на човекот е основа за различни медицински апликации за дијагноза. Но, што ако и нашиот глас може да ни каже многу за моменталната состојба на здравјето на една личност? **Дали гласот (говорот) е нова крв**? Со цел подобро

да ги препознаеме акустичните карактеристики на обрасци на болест што се појавуваат во човечкиот глас, прво мора да го разбереме механизмот на човечкиот производствен глас.

Различни структури во абдоменот, градите, грлото и главата мора да бидат координирани кога правиме звук: дијафрагмата, белите дробови, душникот и градниот кош служат како ветробран. Вокална поддршка, гркланот и глотисот (glottis) како генератор на вибрато и тон, фаринксот (pharynx), усната шуплина и носната шуплина како резонатор. Практично, целото наше тело влијае на звукот на гласот.

Сето ова е предмет на процес на когнитивна контрола заедно со други области на нервниот систем. Кога когнитивниот или физичкиот напор, па дури и болеста, влијае на овој сложен процес, може да генерира карактеристики на јазикот што може да се користи за да се разликува болен од здрав.

Карактеристики се однесуваат на интонација, интензитет и темпо. Ако системи за вештачка интелигенција се обучени со соодветни податоци за звук и говор, кои обично се состојат од двојно аудио и придружна прибелешка, како што е дијагноза, тие можат да препознаат соодветни цели за учење.

.

КОРИСТЕНА ЛИТЕРАТУРА

1. NHS Digital, (2009, September 22). *Autism Spectrum Disorders in Adults Living in Households Throughout England – 2007, Report from the Adult Psychiatric Morbidity Survey*. Retrieved from: http://digital.nhs.uk/catalogue/PUB01131.
2. Bölte, S., and Hallmayer, J. (2011). *Autism Spectrum Conditions: FAQs on Autism, Asperger Syndrome, and Atypical Autism Answered by International Experts*. Boston, MA, Hogrefe Publishing.
3. Cage, E., Di Monaco, J., and Newell, V. (2018). Experiences of Autism Acceptance and Mental Health in Autistic Adults. *J Autism Dev Disord*, 48(2): 473–484.
4. Dawson G. (2008). Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder. *Dev Psychopathol*, 20(3):775-803.
5. Klin, A., Lin, D. J., Gorrindo, P., Ramsay, G., and Jones, W. (2009). Two-year-olds with autism orient to nonsocial contingencies rather than biological motion. *Nature*, May 14; 459(7244): 257–261.
6. Pierce, K., Conant, D., Hazin, R., Stoner, R., and Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Arch Gen Psychiatry*, Jan;68(1):101-9.
7. Klin, A., Jones, W., Schultz, R., Volkmar, F., and Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Arch Gen Psychiatry*, Sep;59(9):809-16.
8. Shic, F., Macari, S., and Chawarska, K. (2014). Speech Disturbs Face Scanning in 6-Month Olds who Develop Autism Spectrum Disorder. *Biol Psychiatry*, Feb 1; 75(3): 10.1016.
9. Chawarska, K, Macari, S., and Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with ASD. *Biol Psychiatry*, Aug 1; 74(3): 195–203.

10. Shic, F., Bradshaw, J., Klin, A., Scassellati, B., and Chawarska, K. (2011). Limited Activity Monitoring in Toddlers with Autism Spectrum Disorder. *Brain Res*, Mar 22; 1380: 246–254.

11. Chawarska, K, Macari, S., and Shic, F. (2012). Context modulates attention to social scenes in toddlers with autism. *J Child Psychol Psychiatry*. Aug; 53(8).

12. Jones, W., Carr, K., and Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Arch Gen Psychiatry*. 65:946

13. Hosozawa, M., Tanaka, K., Shimizu, T., Nakano, T., and Kitazawa, S. (2012). How children with specific language impairment view social situations: an eye tracking study. *Pediatrics*. 129:E1453–E1460.

14. Young, G. S., Merin, N., Rogers, S. J., and Ozonoff, S. (2009). Gaze behavior and affect at 6 months: predicting clinical outcomes and language development in typically developing infants and infants at risk for autism. *Dev Sci*. 12:798–814.

15. Falck-Ytter, T., Bölte, S., and Gredebäck, G. (2013). Eye tracking in early autism research. *Journal of Neurodevelopmental Disorders*. 5:28

16. Elsabbagh, M., Gliga, T., Pickels, A., Hudry, K., Charman, T., Johnson, M., and Team, T. B. (2013). The development of face orienting in infants at-risk for autism. *Behav Brain Res*. 251:147–154.

17. Billeci, L., Narzisi, A., Tonacci, A., Sbriscia-Fioretti, B., Serasini, L., Fulceri, F., Apicella, F., Sicca, F., Calderoni, S., & Muratori, F. (2017). An integrated EEG and eye-tracking approach for the study of responding and initiating joint attention in Autism Spectrum Disorders. *Scientific Reports.* Volume 7, Article number: 13560.

18. Ness, S. L., Manyakov, N. V., Bangerter, A., Lewin, D., Jagannatha, S., Boice, M., Skalkin, A., Dawson, G., Janvier, Y. M., Goodwin, M. S., Hendren, R., Leventhal, B., Shic, F., Cioccia, W., and Pandina, G. (2017). JAKE® Multimodal Data Capture System: Insights from an Observational Study of Autism Spectrum Disorder. *Front Neurosci*. Sep 26;11:517.

19. Loth E, Charman T, Mason L, Tillmann J, Jones EJH, Wooldridge C, Ahmad J, Auyeung B, Brogna C, Ambrosino S, et al. (2017). The EU-AIMS Longitudinal

European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. *Mol Autism*. 8:24.

20. Wade, N. J. (2010). Pioneers of eye movement research. *I-Perception*, *1*(2), 33–68. http://doi.org/10.1068/i0389

21. Bard, C., and Fleury, M. (1976). Analysis of visual search activity during sport problem situations. *J. Hum. Mov. Stud.* 3, 214–222.

22. Ryu D., Abernethy B., Mann D. L., Poolton J. M. (2014). The contributions of central and peripheral vision to expertise in basketball: how blur helps to provide a clearer picture. *J. Exp. Psychol*. 41, 167–185. 10.1037/a0038306

23. Mann D. L., Spratford W., Abernethy B. (2013). The head tracks and gaze predicts: how the world's best batters hit a ball. *PLOS ONE* 8:e58289. 10.1371/journal.pone.0058289

24. Vansteenkiste P., van Hamme D., Veelaert P., Philippaerts R., Cardon G., Lenoir M. (2014). Cycling around a curve: the effect of cycling speed on steering and gaze behavior. PLoS ONE9:e102792. 10.1371/journal.pone.0102792

25. Milazzo N., Farrow D., Ruffault A., Fournier J. F. (2016). Do karate fighters use situational probability information to improve decision-making performance during on-mat tasks? J. Sports Sci.34, 1547–1556. 10.1080/02640414.2015.1122824

26. R. Kredel, C. Vater, A. Klostermann, and E.J. Hossner. (2017). Eye-Tracking Technology and the Dynamics of Natural Gaze Behavior in Sports: A Systematic Review of 40 Years of Research. *Frontiers in psychology, 8*, 1845.

27. Stefanie Hüttermann, Daniel Memmert & Fabian Liesner (2013) Finding the Happy Medium: An Analysis of Gaze Behavior Strategies in a Representative Task Design of Soccer Penalties, Journal of Applied Sport Psychology, 26:2, 172-181, DOI: 10.1080/10413200.2013.816892

28. Castaneda, L., Sidhu, M. K., Azose, J. J., & Swanson, T. (2016). Game Play Differences by Expertise Level in Dota 2, A Complex Multiplayer Video Game. *International Journal of Gaming and Computer-Mediated Simulations (IJGCMS), 8*(4), 1-24. doi:10.4018/IJGCMS.2016100101

29. Díaz, C. M. C., Dorner, B., Hussmann, H., and Strijbos, J-W. (2015). Scientific Heroes: Multiplayer Online Battle Arenas Foster Players' Hypothetico-Deductive

Reasoning. *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*, 481-485, DOI: 10.1145/2793107.2810313

30. Balconi, M., Rezk, S., Leanza, F. (2015). Environment Impact based on functional and aesthetical features: what is their influence on spatial features and rewarding mechanisms?, *Abstract de VI International Conference on Spatial Cognition*, 16 (S1): 83-83.

31. Higuera-Trujillo, J. L., Marín-Morales, J., Rojas, J-C., and Tarruella-Maldonado, J. L. (2016). Emotional maps: neuro architecture and design applications. *IFDP`16 – Systems & Design:Beyond Processes and Thinking.* DOI: 10.4995/IFDP.2016.3170.

32. Ciani, O., et al., (2015). De innovatione: The concept of innovation for medical technologies and its implications for healthcare policy-making. *Health Policy Technol*. 5, 47e64. http://dx.doi.org/10.1016/j.hlpt.2015.10.005.

33. J. Currie, R. R. Bond, P. McCullagh, P. Black, D. D. Finlay and A. Peace. (2018). "Eye Tracking the Visual Attention of Nurses Interpreting Simulated Vital Signs Scenarios: Mining Metrics to Discriminate Between Performance Level", *IEEE Transactions on Human-Machine Systems*, vol. 48, no. 2, pp. 113-124, doi: 10.1109/THMS.2017.2754880.

34. John, K. K., Jensen, J. D., King, A. J., Pokharel, M., and Grossman, D. (2018). Emerging applications of eye-tracking technology in dermatology. *Journal of Dermatological Science*, https://doi.org/10.1016/j.jdermsci.2018.04.002.

35. Balslev T, Jarodzka H, Holmqvist K. et al. (2012). Visual expertise in paediatric neurology. *Eur J Paediatr Neurol*, 16(2):161–166.

36. Benjamin Law, M. Stella Atkins, A. E. Kirkpatrick, and Alan J. Lomax. (2004). Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment. In *Proceedings of the 2004 symposium on Eye tracking research & applications* (ETRA '04). ACM, New York, NY, USA, 41-48. DOI=http://dx.doi.org/10.1145/968363.968370.

37. Garry, Jonah et al. (2016). A pilot study of eye-tracking devices in intensive care. *Surgery*, Volume 159, Issue 3, 938 – 944.

38. White, M., Howes, D., Egan, R., Braund, H., & Szulewski, A. (2017). P128: The novel application of eye-tracking for the cognitive task analysis of expert physician decision-making while leading real-world traumatic resuscitations. *CJEM, 19*(S1), S121-S121. doi:10.1017/cem.2017.330.

39. Harezlak K., Kasprowski P. (2017). Application of eye tracking in medicine: a survey, research issues and challenges. *Computerized Medical Imaging and Graphics*. doi: 10.1016/j.compmedimag.2017.04.006.

40. Tseng P.-H., Cameron I. G. M., Pari G., Reynolds J. N., Munoz D. P., Itti L. (2013). High-throughput classification of clinical populations from natural viewing eye movements. *Journal of Neurology*. 260(1):275–284. doi: 10.1007/s00415-012-6631-2.

41. Laurent Itti. (2015). New Eye-Tracking Techniques May Revolutionize Mental Health Screening. *Neuron*, Volume 88, Issue 3, Pages 442-444, https://doi.org/10.1016/j.neuron.2015.10.033.

42. Duque, A., & Vázquez, C. (2015). Double attention bias for positive and negative emotional faces in clinical depression: evidence from an eye-tracking study. *Journal of behavior therapy and experimental psychiatry, 46*, 107-14.

Прилог 1: Анкетен прашалник

Анкетен прашалник за избор на РИБА-КОНЗЕРВА (со цени)

АНКЕТЕН ПРАШАЛНИК (self-report)

Испитаник:

Во дадениот Тест, **заокружи** ја она риба-конзерва што сакаш да ја купиш:

ТЕСТ



55 денари



55 денари



55 денари



55 денари