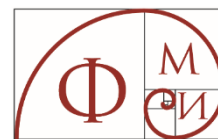




Софийски университет „Св. Климент Охридски“

Факултет по математика и информатика
Катедра "Информационни технологии"



Визуализация на големи обеми от данни

Дафинка Савова Митева

АВТОРЕФЕРАТ

на дисертационен труд

за присъждане на образователна и научна степен "доктор" в
професионално направление "4.6 Информатика и компютърни науки"
докторската програма "Информационни технологии (Информационни
и комуникационни технологии)"

Научен ръководител:
проф. д-р Елиза Стефанова

София, 2023 г.

Съдържание

Благодарности.....	iv
Увод.....	1
Актуалност на проблема	1
Обект.....	2
Предмет.....	2
Цел на дисертационния труд.....	2
Задачи на дисертационния труд.....	2
Структура на дисертационния труд	3
Глава 1: Литературен обзор.....	5
1.1. Учебните аналитики в обучението.....	5
1.2. Визуализация на данни.....	7
1.3. Платформи за създаване на виртуално табло	10
Глава 2: Експерименти с прилагане на различни УА методи и подходи с цел оптимизиране на учебния процес.....	11
2.1. Методология на изследването	11
2.2. Предизвикателства и изводи	11
Глава 3: Експерименти с проектирани, разработени и/или използвани методи за визуализация	12
3.1. Апробирани модели на виртуални табла.....	12
3.2. Сравнение на средствата за виртуално табло	13
3.3. Изводи и препоръки	13
Глава 4: Изисквания към единна интегрирана система за учебни аналитики	14
4.1. Експертното мнение за учебни аналитики	15
4.2. Експертната гледна точка за избор на визуализации на учебните аналитики.....	19
Глава 5: Модел на система за визуализиране на учебни аналитики eAnalytics	21
5.1. Потребители и роли.....	21
5.2. Компоненти на системата	22
5.3. Мобилност, сигурност и етични норми.....	24
Глава 6: Създаване на прототип.....	25
6.1. Избор на средство за реализация на виртуално табло	25
6.2. Данни и метрики, използвани в прототипа	25
6.3. Импортиране на данните и свързването им в модела	26
6.4. Създаване на виртуално табло и избор на визуализации.....	26
6.5. Създаване на шаблон на виртуалното табло	26

6.6. Създаване на приложение и споделяне в онлайн магазин.....	26
Глава 7: Тестване на прототипа.....	26
7.1. Предварителна подготовка на данните.....	27
7.2. Функционално тестване на прототипа.....	27
7.3. Потребителски тестове с протокол „Мисля на глас“	29
Заклучение.....	30
Постигнати резултати.....	30
Насоки за бъдещо развитие.....	31
Авторска справка	32
Приноси на дисертационния труд.....	32
Публикации.....	33
Научни проекти, свързани с дисертационния труд	34
Декларация за оригиналност	35
Библиография	35

Бележка: В автореферата, където не е посочено, всички препратки към страници, глави, приложения, цитирана литература, таблици и фигури, се отнасят до пълния текст на дисертацията.

Благодарности

Конфуций е казал, че който знае всички отговори, не е задал всички въпроси.

Моят научен ръководител проф. д-р Елиза Стефанова ме научи непрекъснато да задавам въпроси и да търся тяхните отговори. Да бъда креативна и да отстоявам целта си докрай. Благодаря за изключителния ѝ професионализъм, за личното отношение, за безценните съвети, винаги дадени в точния момент.

На проф. д-р Красен Стефанов за нестихващата енергичност, с която ме мотивира, и за интересните научноизследователски проекти, чрез които ми даде шанс да представя резултатите от тази дисертация в национални и международни конференции. Благодаря за мъдрите наставления и куража да довърша работата докрай.

На проф. д-р Слави Стоянов, който ме вдъхнови за едно от изследванията в тази дисертация и ми помогна да го реализирам в професионална среда.

На д-р Елица Пелтекова, която с младежкия си ентузиазъм постоянно ме насърчава да не спирам да се боря за всичко, което сама съм избрала да постигна.

На експертите, които споделиха ценния си опит с мен – проф. д-р Елиза Стефанова, проф. д-р Мая Стоянова, проф. д-р Красен Стефанов, проф. д-р Павел Бойчев, доц. д-р Мариана Атанасова, доц. д-р Николина Николова, доц. д-р Александър Димов, доц. Теменужка Зафирова-Малчева, доц. д-р Любка Алексиева, гл. ас. Ирена Авджиева, гл. ас. Мая Желязкова, ас. Албена Антонова, старши преподавател Пенчо Михнев, Диана Пенчева, Йоанна Атанасова, Димитринка Томова, Светлана Димитрова и всички студенти, които с въодушевление приеха предизвикателството да се включат.

На експертите, които, скрити зад анонимното участие, се осмелиха да бъдат по-критични и откровени с мен.

На всички учители, с които бяхме заедно в проучванията и експериментите.

На катедра „Информационни технологии“ във ФМИ на СУ и приятели, които ме карат да се чувствам като част от екип!

На проф. д-р Павел Бойчев за редакторските съвети и полезната обратна връзка.

На моето семейство, което уважи желанието ми да напиша този труд и застана до мен.

Най-искрено благодаря!

Увод

Една картина казва повече от хиляди думи...

Актуалност на проблема

В съвременния свят на информационни технологии и големи данни можем да перифразираме известната поговорка „от много дървета не се вижда гората“ в „от много данни не се виждат резултатите“. Анализирането на големи обеми от данни носи със себе си огромни възможности, но за да могат те да се разгърнат, преди всичко потребителите трябва да съумеят да интерпретират резултатите, да извлекат полезната информация от тях и да я използват. А за да се създаде мост между големите обеми от данни и хората, които вземат решения, тези данни трябва да „говорят“ бързо и ясно, те трябва да бъдат визуализирани.

Тази дисертация изследва как визуализацията на големите обеми от данни може да отговори на предизвикателствата в образованието.

Традиционното обучение е насочено към средностатистическия обучаем, което увеличава риска в даден момент изборът на учебни материали, методите на преподаване, оценката на знанията и резултатите да се окажат предназначени за „несъществуващ“ обучаем. Затова съвременното образование е изправено пред редица предизвикателства и въпроси: Може ли обучението да бъде персонализирано? Може ли всеки обучаем да следва своя индивидуална учебна програма, съобразена със стартовото му ниво, темп на учене, напредъка и специфичните потребности? Могат ли резултатите на обучаемия да бъдат предсказани и усилията да бъдат насочени там, където среща затруднения? Може ли чрез отговор на личните потребности да бъде предизвикан и запазен интересът на обучаемите? Как може обучението да бъде по-ефективно и полезно? Как да се направи оценка в каква посока да се развият/подобряват учебните материали и курсове?

Отговор на тези въпроси могат да дадат учебните аналитики (Learning Analytics, LA, UA), дефинирани официално през 2011 година, като „измерване, събиране, анализ и отчитане на данни за обучаемите и контекста, в който се използват, за целите на разбирането и оптимизирането на обучението и средата, в която то се осъществява“¹.

Според Siemens (Siemens et al., 2011) използването на UA може (1) да намали броя на отпаднали обучаеми чрез ранното откриване на учащи в риск и уведомяване на преподавателите, (2) да персонализира учебния процес и съдържание, така че всеки обучаем да получава ресурси според текущите си познания, (3) своевременно да уведомява преподавателите кои обучаеми се нуждаят от допълнителна помощ и кои практики на преподаване оказват най-голямо въздействие, (4) да допринесе за създаване на висококачествен дизайн на обучението, (5) да създаде интерактивни визуализации на сложна информация, които ще помогнат за бързото намиране на отговори и др.

¹ 1st International Conference Learning Analytics and Knowledge, <http://www.wikicfp.com/cfp/servlet/event.showcfp?eventid=11606>

Съвременното обучение използва софтуерни платформи за управление на учебното съдържание, които осигуряват технологичен апарат паралелно с публикуването на учебното съдържание да се събира и съхранява информация и за активността на потребителите – преподаватели и обучаеми. В рамките на една образователна организация могат да се използват повече от една системи, обслужващи различни звена, специфични курсове, работа по конкретни проекти и задачи. Понякога естеството на обучението изисква използване на технологии и средства, които поддържат своя специфична среда или изискват конкретна такава. Често цели курсове се провеждат в платформи, извън административния контрол на основната образователна институция. За да получат меродавен отговор поставените в началото въпроси, учебните аналитики от всички системи, обслужващи обучението, трябва да бъдат обединени и анализирани като едно цяло.

В преглед на литературните източници за периода 2000 – 2017, публикувани в специализирани издания за визуализация и учебни аналитики (Vieira et al., 2018), авторите заключават, че (1) има сравнително малко инструменти за визуален анализ на учебни аналитики, (2) за визуализация на учебни аналитики все още се използват предимно традиционни техники за представяне на статистически данни, като лентови и точкови диаграми и (3) липсват проучвания, които да навлизат по-дълбоко в образователните теории и в същото време да прилагат по-сложни съвременни визуализации.

Обект

Обект на изследването, описано в този дисертационен труд, са учебните аналитики, използвани в системите за управление на обучението, и тяхната визуализация.

Предмет

Предмет на изследването са видовете справки, които една система за управление на обучението може да предостави в частта за учебни аналитики и методите за визуализация на данните за по-лесното им и правилно анализиране.

Цел на дисертационния труд

Целта на това проучване е проектиране на система за учебни аналитики, в която чрез съвременни методи за визуализация да се очертаят начини за повишаване ефективността на електронното обучение и да се подпомогнат обучаемите, преподавателите и мениджърите на учебни заведения при вземането на правилни решения в точния момент.

Задачи на дисертационния труд

За постигане на поставената цел са решени следните задачи:

Задача 1. Да се изследват и анализират предизвикателствата пред създаването на система за визуализиране на учебни аналитики, в следните аспекти:

- 1.1. метрики и ключови индикатори за ефективност, методите за събиране на данни за обучаеми, моделите за прилагане на учебни аналитики, спазване на етичните норми и защитата на личните данни.
- 1.2. съвременните методи за визуализиране на данни и големи масиви от данни.
- 1.3. съвременните платформи за създаване на виртуални табла

Задача 2. Да се изследва и анализира как методите и средствата за УА могат да оптимизират учебния процес и да повишат успеваемостта на обучаемите.

Задача 3 Да се изследват и анализират функционалностите на виртуалното табло като средство за визуализация на учебни аналитики в системите за управление на обучението (СУО).

Задача 4. Да се проучат и систематизират очакванията на потребителите на СУО по отношение на:

- 4.1. справките и отчетите, които се надяват да открият и използват;
- 4.2. справките, които различните категории потребители (обучаеми, преподаватели и мениджъри) намират за най-интуитивни и лесни за разчитане.

Задача 5 Да се създаде модел на система за визуализиране на учебни аналитики.

Задача 6 Да се създаде прототип на система за учебни аналитики.

Задача 7 Да се тества и анализира създадения прототип.

Структура на дисертационния труд

Дисертационният труд се състои от увод, седем глави, заключение, авторска справка, библиография и шест приложения.

Уводът представя актуалността на проблема, дефинира обект, предмет и цел на дисертационния труд, описва задачите и структурата.

В **глава 1** е направен литературен обзор на различните видове учебни аналитики и тяхното използване в образованието. Посочени са метриките и ключовите показатели за ефективност. Изследвани са методите за събиране на данни в СУО. Коментирани са етичните норми и изискванията за защита на личните данни. Направен е преглед на съвременните методи и средства за визуализация на учебни аналитики. Анализирани са съвременните платформи за създаване на виртуално табло.

В **глава 2** са описани експерименти, проведени в рамките на дисертацията, разглеждащи как прилагането на различни учебни аналитики може да оптимизира учебния процес, като подпомогне преподаватели и мениджъри на учебни заведения, а оттам и самите обучаеми.

В **глава 3** са описани експерименти, апробиращи различни методи за визуализация, създадени и използвани по различни научни проекти в рамките на дисертацията.

В **глава 4** са представени изискванията за създаване на интегрирана система за учебни аналитики. Описани са две проучвания, проведени в рамките на дисертацията. Първото проучване представя експертното мнение за използването на големи обеми от данни в полза на обучението чрез създаване на учебни

аналитики в системите за управление на обучението. Второто проучване разкрива визията на експертите за това кои визуализации позволяват най-бързо и точно тези големи обеми от данни да бъдат разчетени и възприети.

В **глава 5** е предложен архитектурен модел на система за визуализиране на учебни аналитики. Неговата цел е да отговори на съвременните изисквания за създаване на интегрирана система за учебни аналитики, които да подпомогнат управлението на обучението чрез по-добра визуализация.

В **глава 6** е описано създаването на прототип на модела. Мотивиран е изборът на средство за реализация на виртуално табло. Описани са данните и метриците, които се визуализират в прототипа. Представено е създаването на готово виртуално табло и на шаблон за виртуално табло.

В **глава 7** е описано експерименталното апробиране на модела и прототипа. Приложени са два метода за тестване: функционално тестване и „мисля на глас“ (Think aloud). Описана е методологията на анонимизиране на реалните данни и са онагледени получените резултати.

В **заклучението** е представено обобщение на дисертационния труд с изводи и насоки за бъдещо развитие.

В **авторската справка** за представени приносите на дисертационния труд, описани са публикациите на автора, свързани с темата, и научните проекти, в рамките на които авторът е провел научните експерименти и изследвания, свързани с дисертационния труд

В **Библиографията** са цитирани използваните източници.

Приложение 1: Писмо-покана за участие в задание „Мозъчна атака“ (Brainstorming) за генериране идеи какви справки в СУО очаква да намери съвременният потребител.

Приложение 2: Писмо-покана за участие в дейност „Групиране“ и „Сортиране“ на направените предложенията за справки в СУО.

Приложение 3: Социални въпроси към участниците в дейност „Сортиране и градиране“

Приложение 4: Предложения за справки в частта за учебни аналитики на СУО, предложени от експертите по време на груповата концептуална карта

Приложение 5: Избор на финален брой клъстери при обработката на предложения с метода групово концептуална карта

Приложение 6: Справки по категории, получили над средната оценка по двата критерия: „полезност“ и „приложимост“

Глава 1: Литературен обзор

В тази глава е описана работата по първата задача, като са изследвани и анализирани предизвикателствата пред създаването на система за визуализиране на учебни аналитики.

1.1. Учебните аналитики в обучението

Тази секция описва проучването и анализа на метриците и ключовите индикатори за ефективност, измерващи напредъка на обучаемите, методите за събиране на данни за обучаеми, моделите за прилагане на учебни аналитики, спазване на етичните норми и защитата на личните данни.

1.1.1. Еволюция на УА и тяхното въздействие върху обучението

През 2010 г. George Siemens представя учебните аналитики като "използване на интелигентни данни, произведени от обучаемите, и модели за анализ, за да се открие информация и социални връзки, както и да се предвиди и консултира обучението" (Siemens, 2010). Една година по-късно по време на първата международна конференция за анализ на обучението и знанието (Learning Analytics and Knowledge – LAK) е решено, че учебните аналитики ще се фокусират върху интегрирането на технологичните и социално-педагогическите измерения и е дадено съвременното определение, публикувано на сайта на конференцията.

Бързото развитие на информационните технологии през 21-ви век променя единиците и мащабите за измерване на информация, както и методите за обработка на данни. Започвайки с байтове и килобайти преди пет десетилетия, днес говорим за ексабайти (EB, 10¹⁸), зетабайти (ZB, 10²¹) и дори йотабайти (YB, 10²⁴) данни. Събират се големи масиви от данни, които "надхвърлят възможностите на обичайно използваните хардуерни среди и софтуерни инструменти за улавянето, управлението и обработката им в приемлив за потребителите период от време"².

През 2012 г. Gartner дава дефиниция на големите данни, известна още като трите „V“-та: *volume* (какъв е обемът на данните), *velocity* (колко бързо се обработват тези данни), *variety* (колко разнообразни са по тип). Към тази дефиниция някои автори добавят още "V"-та, описващи модела на големите данни от бизнес гледна точка (Mediratta, 2015): *variability* – отчита дали потокът от данни е редовен и надежден, *veracity* – проверка дали данните са достоверни, преди да се извърши какъвто и да е анализ и *value* – идентифицира целта за обработка на данните, анализът трябва да се придържа към етичните норми, да пази репутацията на потребителите. Wang (R. Wang, 2012) добавя *viscosity*, с което измерва „съпротивлението“ на данните, когато пристигат от различни източници и с различна скорост и *virality*, с което описва колко бързо тази информация се разпространява сред потребителите в мрежи от тип "peer-to-peer" (с равноправни участници).

² Big Data Definition, Teradata, <https://www.teradata.com/Glossary/What-is-Big-Data>, последен достъп 01.02.2023

Обработката на данни, генерирани в учебните среди, постепенно се превръща в истинско предизвикателство, наложило използването на технологии за големи данни и инструменти за тяхната обработката. В преглед на най-съвременните инструменти за извличане на данни и анализ на обучението във висшето образование се посочват техниките и инструментите за обработка на големи данни, които могат да се използват по различни начини в учебните аналитики (Salihoun, 2020).

1.1.2. Метрики и ключови индикатори за ефективност

Учебните аналитики използват метрики и ключови индикатори за ефективност (KPI)³, с които измерват и анализират напредъка на обучаемите.

В преглед на литературата (Schwendimann et al., 2017) авторите систематизират над 200 индикатора, които се използват в съвременните системи за обучение. Групирани в 6 категории, те измерват: (1) *Действия на обучаемия* – брой на посетените страници; брой на изтеглените файлове; време, прекарано в системата; (2) *Качества на обучаемия* – предишно обучение и курсове, добити компетенции, възраст; (3) *Отношение към съдържанието на учебния материал* – завършени и пропуснати теми, уловени настроения от съобщенията във форумите; (4) *Постигнати резултати* – среден успех, разпределение на резултатите в групата; (5) *Контекст на обучението* – в какъв контекст се провежда обучението, къде се провежда и (6) *Социален аспект* – комуникации между участниците в групата.

Според (Hout, N., 2020) може да се обобщи, че критичните индикатори в повечето съвременни СУО се свеждат до: (1) данни за участие в курс, (2) брой на обучаемите, успешно завършили и брой на отпадналите, (3) детайли за активността на обучаемите, (4) най-често преглежданите учебни материали, (5) данни за резултатите, (6) времето, прекарано в разглеждане на учебните материали, (7) обучаемите с най-добро и най-слабо представяне, (8) търсене на модели в поведението на обучаемите, (9) опити за изпълнение на задания и (10) къде физически се намират обучаемите.

1.1.3. Методи за събиране на данни

Учебните аналитики предполагат използването на широк набор от методи за анализ на данни, базирани на познания в различни сфери, като статистика, психология, компютърни науки, образование (Khalil & Ebner, 2016). Тези методи са разгледани и описани в детайли в редица статии. Като цяло подходите се групират в две категории: анализ на данните, генерирани от потребителите при натискане на бутоните на мишката, докато обхождат веб-страници (click-stream) и обработка на естествен език (Natural Language Processing, NLP).

1.1.4. Модели за прилагане на учебни аналитики

Големите обеми от данни, които се събират в електронните системи за обучение, трябва да бъдат управлявани, за да може от тях да се извлича полезната информация, която да подпомага вземането на решения.

³ Key Performance Indicators, <https://www.kpi.org/kpi-basics>, последен достъп 01.02.2023

В пълната версия на дисертацията са разгледани следните модели за прилагане на УА: (1) модел за жизнения цикъл на УА на Clow; (2) модел за прилагане на УА на Elias; (3) референтен модел на Chatti за прилагане на УА; (4) модел на Freitas и (5) модел на Siemens.

1.1.5. Етични норми и защита на личните данни

Успоредно с разрастването на технологичните възможности за анализ и извличане на информация от данните за обучаеми назрява и етичният въпрос „трябва ли“, след като „можем“. Водят се много дебати относно етиката при работа с големи данни и изкуствен интелект, засягащи неприкосновеността на личния живот, опасността от погрешно прогнозиране на нечие поведение и др. Тези потенциални рискове важат и за сферата на образованието.

Съхранението на данните в облачно пространство и използване на съвременни механизми за обработка в мрежова среда налагат създаване на общ правен инструмент за защита на данните, който влезе в сила на 25 май 2018 за Европейския съюз и е известен като General Data Protection Regulation (GDPR). Той регламентира обработката на лични данни, разграничава публични и частни лица и операции, предвижда високи санкции за нарушения. От етична гледна точка събирането и обработката на големи обеми от данни трябва да намери баланс между обхвата на данните, полезността и рисковете от предоставянето на достъп (Florea & Florea, 2020).

1.2. Визуализация на данни

Тази секция описва работата по втория аспект на първата задача – да се изследват и анализират съвременните методи за визуализиране на данни и големи масиви от данни. Целта на този анализ е на по-късен етап да подпомогне избора на визуализации в проектираната система за онагледяване на учебни аналитки.

1.2.1. Силата на визуализацията

В редица научни изследвания психолозите доказват експериментално, че човешкият мозък възприема по-лесно и запаметява по-трайно визуалната информация, отколкото словесната (Whitehouse et al., 2006).

Използването на визуализация провокира вградения в мозъка ни „софтуер“ за откриване на шаблони, корелации и тенденции (Iliinsky & Steele, 2011), които трудно биха били забелязани в традиционните текстове и таблици. Тя може да разкрие нови предизвикателства и да вдъхнови за по-задълбочен анализ.

1.2.2. Методи за визуализация

През 2010 г. Abela описва алгоритъм в 10 стъпки за визуално въздействие върху потребителя (Abela, 2010). Той предлага йерархична диаграма, подобна на блок-схема, която чрез водещи въпроси и отговори, свързани с типа на наличните данни, подпомага потребителя да избере подходяща визуализация. На първо ниво данните се класифицират в 4 основни категории: сравнение, композиция,

разпределение или отношение. На всяко следващо ниво съответната категория се разделя на под-категории.

През 2015 Edo Van Dijk⁴ добавя още два принципа към модела на Abela – отклонение и тенденция. Целта на отклонението е да покаже кои стойности се отклоняват от нормата чрез колонна, стълбовидна, линейна и площна диаграми, докато *тенденцията* показва очакваното развитие на някои променливи чрез линейна, колонна, точкова или площна диаграми.

Методи за визуализиране на различни типове данни

В тази секция са описани примерни методи за визуализиране на данни, които могат да подпомогнат представянето на УА (Митева & Стефанова, 2018).

Например за *сравнение на данни* се използват колонни и стълбовидни, кръгови и пръстеновидни, точкови и площни диаграми, а за *разпределение на стойности* в даден диапазон – хистограма. *Отношения между обекти* могат да се покажат чрез понятийна карта или диаграма на Вен. За представяне на *поток от данни* са подходящи диаграмите от типа блок-схема. За визуализиране на *многомерни данни* се използват диаграми тип радар. За онагледяване на *конкретна числова стойност* обикновено се използва измервателен уред. При *семантичен анализ* за визуализиране на честотата на използваните думи най-подходящ е словесният облак (wordcloud). За представяне на *йерархични данни* могат да се използват интерактивна топлинна и йерархична карта или диаграма-слънце. За визуализиране на *поток от данни* е полезна поточна диаграма (streamgraph), за онагледяване на *взаимовръзки между обекти* - диаграма на връзките (chord diagrams).

Методи за статична и интерактивна визуализация

Статичната визуализация се фокусира върху конкретно значение на данните и може да покаже тяхната стойност в даден момент. Интерактивната визуализация дава възможност на потребителя сам да управлява как ще разглежда данните. Ако стандартната визуализация показва нещата „какви са“, то интерактивната се опитва да даде отговор на въпроса „какво би станало, ако“ (Rouse et al., 2017). Съществуват редица интерактивни техники за взаимодействие между потребителя и графичната визуализация, някои от които са описани от Khan (Khan & Khan, 2011):

- увеличаване и намаляване (zoom in, zoom out), което позволява на потребителя да детайлизира в различна степен видимостта на обектите;
- общ преглед и детайлизиране, показващ едновременно местоположението на обекта и подробен изглед на избраната област;
- фокусиране и детайлизиране на контекст (fisheye view), което позволява концентриране върху конкретна област и подробния ѝ преглед, без тя да бъде изваждана от контекста на цялото изображение.

⁴ Which chart type works best for your data?, Blog post, <http://www.chartlr.com>, последен достъп: 21.05.2022

Методи за визуализиране на големи обеми от данни

Електронните среди за обучение позволяват събирането на големи обеми от данни. Често тези данни идват от различни източници и са представени в структуриран или неструктуриран формат. Към тях се добавят и данните, генерирани от средствата за социална комуникация в курса, които се генерират в реално време и с голяма скорост. Съвременните методи за визуализация трябва да могат да представят ефективно динамични обеми от данни.

Bikakis (Bikakis, 2018) описва най-съвременните подходи за управление на данните, извличане на знания и визуализация на информацията. Визуализирането на големи обеми от данни ни изправя пред редица предизвикателства, описани от Wang (L. Wang et al., 2015), като загуба на информация при намаляване на данните, появата на визуален шум, когато данните са много близки, трудното възприемане на големи изображения, които се променят с голяма скорост, изискванията за висока производителност на компютърната техника, с която се обработват.

Съвременните системи за електронно обучение събират, съхраняват и обработват на практика данни за всичко, което се случва в образователния процес. Можем да получим подробна информация за всеки аспект от обучението, но най-голямото предимство на визуализацията е, че всичко може да се покаже на един екран като моментна снимка на данните. Такъв визуален инструмент е виртуалното табло за управление (dashboard).

1.2.3. Средства за визуализация на учебни аналитики

В контекста на тази дисертация средствата за учебни аналитики са разглеждани и оценявани от гледна точка на предложените в тях визуализации.

В ревю на средствата за учебни аналитики във висшето образование авторът (Salihoun, 2020) разглежда около 40 инструмента, често използвани за събиране на данни и анализи в областта на образованието.

С широкото навлизане и популярността на MOOC курсовете, на дневен ред излиза необходимостта от аналитики, изучаващи поведението на обучаемия при учене чрез използване на данни от натискане бутоните на мишката при разглеждане на видеа. Примерна визуална аналитична система е VisMOOC, описана в (Shi et al., 2014). На по-късен етап един от авторите на VisMOOC представя две нови средства за визуализация на MOOC курсове (Chen, 2018): PeakVizor и ViSeq. В проучване (Caldarola & Rinaldi, 2017) авторите описват 36 от най-използваните инструменти и техники за визуализирането на големи обеми от данни, групирани в 4 категории: (1) за визуализиране на информация; (2) за визуализиране на данни; (3) за визуализиране на научни изследвания и (4) инструменти за бизнес интелигентност и визуализация.

Виртуално табло

Stephen Few дефинира виртуалното табло като „визуално показване на най-важната информация, необходима за постигане на една или повече цели; консолидирани и подредени на един екран, така че информацията да може да се наблюдава с един поглед” (Few, 2006). Първоначалната идея идва от таблото на

превозни средства или самолети, което обобщава състоянието на всички двигателни системи в малък панел. Pauwels (Pauwels et al., 2009) обобщава целите на виртуалното табло като (1) налагане на последователност в мерките и процедурите за измерване; (2) мониторинг за ефективност и ранно прогнозиране на потенциални проблеми; (3) планиране на бъдещи цели и стратегии и (4) комуникация с важни заинтересовани страни, като оценяването става чрез избор на метрики/показатели.

Таблото за управление може да се разработи като самостоятелно приложение или уеб-базиран инструмент за анализ като част от СУО. Дизайнът на приложението и показателите за ефективност зависят от конкретните цели.

1.3. Платформи за създаване на виртуално табло

Тази секция описва работата по третия аспект на първата задача – да се изследват и анализират съвременните платформи за създаване на виртуално табло. Целта на този анализ е да подпомогне създаването на интерактивно виртуално в проектираната система за визуализиране на УА.

През 2022 водещата изследователска компания Gartner⁵ публикува доклад, в който оценява 20 доставчици на платформи за анализ и бизнес интелигентност (BI). Към този тип платформи, съгласно дефиницията на Gartner, спадат онези, които осигуряват леснотата и удобството за потребителя да „анализира, изследва, споделя и управлява данни“, дават му възможност да визуализира своите и открива чужди прозрения, често базирайки се на изкуствен интелект. В това изследване платформите са оценявани по критериите: (1) сигурност; (2) управление; (3) облачен анализ, (4) свързване на източници на данни; (5) подготовка на данни; (6) каталог; (7) автоматизирани прозрения; (8) визуализация на данни; (9) запитване на естествен език; (10) разказване на данни, (11) генериране на естествен език и (12) споделяне. Оценката се основава на всички публикувани ревюта, като по-старите оценки се намаляват наполовина на всеки 12 месеца.

На тяхна база Gartner дефинира своя „магически квадрант“, в който компаниите са категоризирани като „лидери“, „претенденти“, „визионери“ и „нишови играчи“. За 2022 г. лидерската позиция в съчетаване на визия и предоставяне на стойност заема Майкрософт. Power BI, в съчетание с Azure и Офис 365, предлагат визуално базирано представяне на данни, разширен анализ и интерактивни виртуални табла за управление. Следващата позиция заемат Tableau и Qlik. С малко над средното ниво по пълнота на визията, но над средното ниво по способност за изпълнимост се нареждат продуктите на Google и Domo.

В пълната версия на дисертацията са разгледани предимствата и предизвикателствата пред всяка от платформите Microsoft Power BI, Tableau, Qlik, Google Data Studio и Domo и е направено заключение, че Microsoft Power BI позволява свързването на най-много типове източници на данни от гледна точка на образованието, позволява обработка и локално съхранение, изключително

⁵ Gartner Research, <https://www.gartner.com/en/documents/4011116>, последен достъп: 01.02.2023

богат набор от визуализации, както и възможности да бъдат персонализирани наличните и програмирани нови такива, предлага лесен и интуитивен за потребителя интерфейс, поддържа мобилна версия, може да бъде използван безплатно за създаване на визуализации, а цената за споделяне не е висока.

Глава 2: Експерименти с прилагане на различни УА методи и подходи с цел оптимизиране на учебния процес

В тази глава е описана работата по втората поставена задача – да се изследва и анализира как методите и средствата за УА могат да оптимизират учебния процес и да повишат успеваемостта на обучаемите.

В рамките на тази дисертация е проведено изследване как методите и средствата за УА повишават успеваемостта на обучаемите (Miteva et al., 2017). Изследователският метод е фокусиран върху 3 основни цели: (1) подобряване качеството на обучение за студентите; (2) подобряване на качеството на преподаване от преподаватели; (3) подобряване на цялостната организация на провеждане на курс на ниво факултет.

2.1. Методология на изследването

Изследването се базира на три електронни курса „Въведение в компютърните мрежи“, „Основи на маршрутизирането и комутирането“ и „Мащабируеми мрежи“, проведени през учебната 2015/2016 година с 98 първокурсници от магистърски програми „Разпределени системи и мобилни технологии“ и „Защита на информацията в компютърните системи и мрежи“ при Факултета по математика и информатика на Софийски университет (СУ).

За по-детайлен анализ данни за поведението на обучаемите са събирани от всички възможни източници, свързани с процеса на обучение. От вътрешна за университета система за управление на студентската информация (СУСИ) са взети организационни данни, от СУО на СУ (Moodle) – данни за активността на студентите и от системата с учебни материали и изпитни тестове Instructure Canvas данни за резултатите и активността. Всички събрани данни са съхранени в централизирана база от данни (УА база от данни) и допълнително са анализирани с помощта на стандартни статистически софтуерни инструменти – в конкретния случай езика R и средата R Studio. Накрая тази база с резултати е анализирана от образователни експерти, които са формулирали списък с препоръки за подобряване на процеса на обучение от гледна точка на всички заинтересовани страни.

2.2. Предизвикателства и изводи

Изграждането на една система за учебни аналитики се изправя пред много предизвикателства. Подобна система трябва да получи достъп до възможно най-голям набор от данни. Едно примерно предизвикателство е да предложи автоматичен анализ, базиран на данни от повече от един курс. СУО работят на ниво

курс, вградените журнали показват статистически данни в рамките на един курс. Когато искаме да проследим как учебната програма се възприема от обучаемите и как те се справят с учебния материал, е необходим по-комплексен анализ.

Технологично предизвикателство е да се събират, съхраняват и обработват данни, импортирани от външни системи, без да има предварително дефинирана релационна схема, определяща как да се съпоставят таблици с данни. Импортирането на данни от учебна система, която е извън административния контрол на институцията и извън правомощията да бъде регистриран потребителски акаунт със специални права за извличане на данни.

Предизвикателство е и езиковата разлика. Системите могат да съхраняват данните както на български език, така и на английски език. На обучаемите се дава право да редактират профилите си, да променят лични данни като имена и имейл адреси. Трябва да бъде намерен начин как тези данните да бъдат съпоставени.

В тази глава е описан експериментален анализ на прилагането на различни методи и подходи за УА за оптимизиране на учебния процес за подобряване качеството на обучение за студентите; повишаване на качеството на преподаване от страна на преподавателите и подобряване на цялостната организация на обучението на ниво учебна институция.

Глава 3: Експерименти с проектирани, разработени и/или използвани методи за визуализация

В тази глава е описана работата по третата поставена задача – да се изследват и анализират функционалностите на виртуалното табло като средство за визуализация на учебни аналитики в СУО. Резултатите от този анализ впоследствие са използвани при проектирането на виртуално табло при изграждането на целевата система за визуализиране на УА.

3.1. Апробирани модели на виртуални табла

В рамките на тази дисертация са апробирани няколко модела на виртуални табла, създадени и използвани по различни научни проекти. В тази глава е описано сравнение между тях, като са направени изводи и препоръки:

- Moodle - една от най-популярните и широко разпространени СУО с отворен код. За всеки курс се предоставя набор от отчети, които филтрират по заявка събраните данни в log файловете на системата и извличат информация.
- Graphical Interactive Student Monitoring Tool (GISMO) е добавка към Moodle и предоставя цялостна система за наблюдение и проследяване на активността на обучаемите в даден онлайн курс в СУО, като посещаемост на курсове, преглед на материали, предаване на задания и др.
- Canvas Analytics – апробирана в периода от 2012 до 2019 г. Сиско мрежовата академия на СУ “Св. Климент Охридски”, част от Cisco Networking Academy

- Виртуално табло на системата Learning Analytics and Reflection & Awareness Environment (LARAe) - създадено в рамките на проект „Working Environment with Social and Personal Open Tools for inquiry based learning (weSPOT)“ за обслужване на учебните аналитики при проиграването на учебни сценарии с използване на изследователския подход в обучението.
- Mobile Inquiry Coordinator Interface⁶ - създаден за наблюдение и координация на обучаемите в мобилните дейности, докато прилагат изследователския подход в обучението.

3.2. Сравнение на средствата за виртуално табло

В Таблица 1 са систематизирани ключовите характеристики на описаните средства за създаване на виртуално табло. От сравнението им може да се заключи, че нито един от изследваните инструменти за създаване на виртуално табло към СУО не отговаря на всички необходими критерии.

Таблица 1: Сравнение на инструментите за виртуално табло

Критерий	Moodle	GISMO	Canvas	LARAe	MICI
Визуализиране на оценките	✓	✓	✗	✗	✗
Графична визуализация на данните	✗	✓	✓	✓	✓
Удобен потребителски интерфейс за дефиниране на заявки	✗	✓	✓	✓	✓
Възможност за дефиниране на персонална заявка, извън вградените	частично	✗	✗	✗	✗
Мобилна версия	✗	✗	✗	✗	✓
Актуализация на версиите	✓	✗	✓	✗	✗

3.3. Изводи и препоръки

Описаните инструменти за УА предоставят голям набор от предварително дефинирани аналитични отчети като активност на студентите, участието им в курсове, статистически данни от регистрационните файлове на СУО и др. Тази информация е изключително ценна, но не е достатъчна, за да отговори на всички въпроси, които могат да възникнат по време на преподаването и ученето. Предложените справки не са достатъчно гъвкави и не могат да осигурят желаната свобода на участниците в процеса на обучение. Често данните изискват допълнителна обработка с външен софтуер за извличане на полезна информация и знания за преподаването и ученето, което означава повече математическа, статистическа и компютърна грамотност, повече работно време, повече софтуерни лицензи.

Едно предизвикателство пред съвременната система за визуализиране на учебни аналитики е да предложи **ориентиран към потребителя интерфейс**, позволяващ лесно и бързо съставяне на заявки.

⁶ weSPOT MICI приложение, <https://apkcombo.com/wespot-mici/els.wespot>, последен достъп: 01.02.2023

Представените инструменти са предвидени да работят със структурирани бази от данни. За целите, за които са били разработени, за количеството и типовете данни, които е трябвало да обработват, тези инструменти са изпълнили своята задача. Един съвременен инструмент за обработка на големи масиви от данни обаче трябва да може **да обработва и неструктурирани данни в реално време**. Подобни решения се използват успешно в бизнеса. Например Hadoop е основната файлова система и инструмент за проследяване на задачи в Amazon Web Services (<https://aws.amazon.com>), Cloudera (<https://cloudera.com>) и др.

Описаните инструменти, както и самите СУО, са ориентирани към конкретен курс или изследване (при инструментите за обслужване на изследователския подход в обучението). Преподавателят може да проследява дейностите на всеки обучаем в рамките на курс, но не може да сравнява представянето на този обучаем в различни курсове от учебната програма и да прави по-генерални изводи.

Един съвременен инструмент за УА трябва да предлага и **междукурсови анализи**. Този тип статистики могат да се използват за подобряване на качеството на образованието чрез овластяване на преподавателя да прави самоанализ; чрез търсене на причините за контраста в резултатите на обучаемите от различни провеждания на един и същи курс или от различни курсове; чрез оценка на въздействието на методите на обучение върху крайните резултати.

От всички причини, споменати по-горе, може да се заключи, че е необходимо да се проектират и разработят нови технологии по отношение на подходи, методологии и софтуерни инструменти, които да са в състояние да визуализират сложната взаимовръзка между обучаеми, преподаватели и курсове. Такъв инструмент може да бъде вградена единица в СУО, но извличането на статистика за различни курсове би влязло в противоречие с ориентираната към курс идеология на СУО. Затова препоръката е да бъде разработено като самостоятелно приложение.

Глава 4: Изисквания към единна интегрирана система за учебни аналитики

В тази глава е описана работата по третата поставена задача – да се изследват и анализират функционалностите на виртуалното табло като средство за визуализация на учебни аналитики в СУО. Резултатите от този анализ впоследствие са използвани при проектирането на виртуално табло при изграждането на целевата система за визуализиране на УА.

Първото проучване представя експертното мнение за използването на големи обеми от данни в полза на обучението чрез създаване на учебни аналитики в СУО. Целта му е да извлече основните изисквания, върху които да бъдат концентрирани усилията в разработката на системата.

Второто проучване разкрива визията на експертите за това кои визуализации позволяват най-бързо и точно тези големи обеми от данни да бъдат

разчетени и възприети. На резултатите от него е базирана реализацията на визуализациите на учебни аналитики, избрани от експертите в първото проучване.

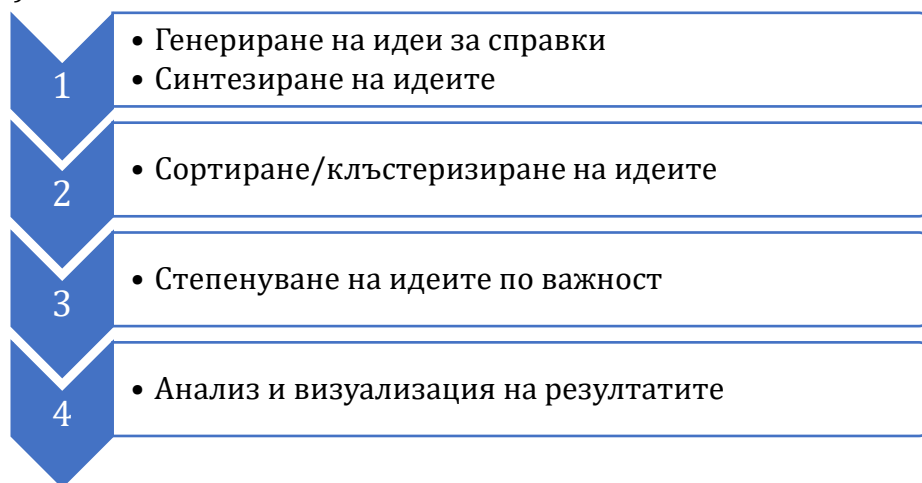
4.1. Експертното мнение за учебни аналитики

В резултат на първото проучване (Miteva & Stefanova, 2020a) са обобщени най-ценните справки и отчети, които потребителите на СУО очакват да намерят в частта за учебни аналитики. Резултатите показват, че експертите с богат опит имат по-високи изисквания и очаквания. Дори за съществуващи справки експертите предлагат подобрени варианти и добавяне на детайли. Благодарение на критичната оценка са извлечени нови изисквания към LA. Получените резултати са използвани за проектирането на система за LA, която чрез съвременни методи за визуализация подпомага студентите, преподавателите и мениджърите на учебни заведения при вземането на правилни решения в точния момент.

4.1.1. Методология на изследването

За събиране на експертното мнение и анализ на данните е използван методът *групова концептуална карта* (Group Concept Mapping) (Kane & Rosas, 2017). Използвана е специализираната онлайн среда на Concept Systems Inc. Global MAX.

Изследването е проведено в следната последователност от дейности (Фигура 1):



Фигура 1: Дейности по извличане на справките за СУО

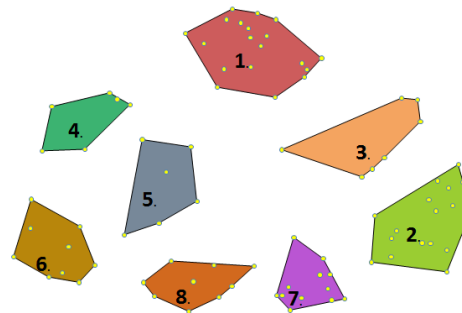
В проучването се включват 30 опитни специалисти от Факултета по математика и информатика, Факултета по педагогика и Центъра по технологии на информационното общество на СУ „Св. Климент Охридски“, филолози, учители по природни науки, докторанти и студенти. Ключовият въпрос, в отговор на който експертите трябва да генерират идеи по време на първата фаза, е *„В раздела за учебни аналитики на Системите за управление на обучението (СУО, например Moodle) бих искал/а да има справки за...“*. Като резултат са синтезирани 85 експертни предложения за справки за частта за учебни аналитики на СУО

4.1.2. Анализ на данните

За обработка на данните в Global MAX се използват статистическите методи многомерно скалиране (multidimensional scaling) и йерархичен клъстерен анализ

(hierarchical cluster analysis). Резултатите от сортирането се представят чрез корелационна матрица, наречена матрица на сходството (similarity matrix), в която за всеки две предложения се отбелязва 1 – ако са сортирани в една и съща група и 0 – ако са разпределени в различни категории. По метода на многомерното скалиране тази матрица се визуализира като *карта на точките*, в която всяка идея се представя като точка в равнината. В следващата фаза от анализа на данните идеите са групирани в категории (кълъстери) по метода на йерархичния кълъстерен анализ. В резултат са получени 8 кълъстери, именувани съгласно основната идея, която обединява предложенията в тях. На Фигура 2 е показан списъкът на категориите, като в скоби за всяка от тях е посочено колко предложения за справки в СУО са включени в нея.

1. Оценяване на преподаватели (19)
2. Оценяване на обучаеми (17)
3. Резултати (7)
4. Обратна връзка за курс (6)
5. Активност в системата (6)
6. Подпомагане на обучаемите (9)
7. Активност на обучаемия (12)
8. Управление на курс (9)



Фигура 2: Финален брой категории със справки

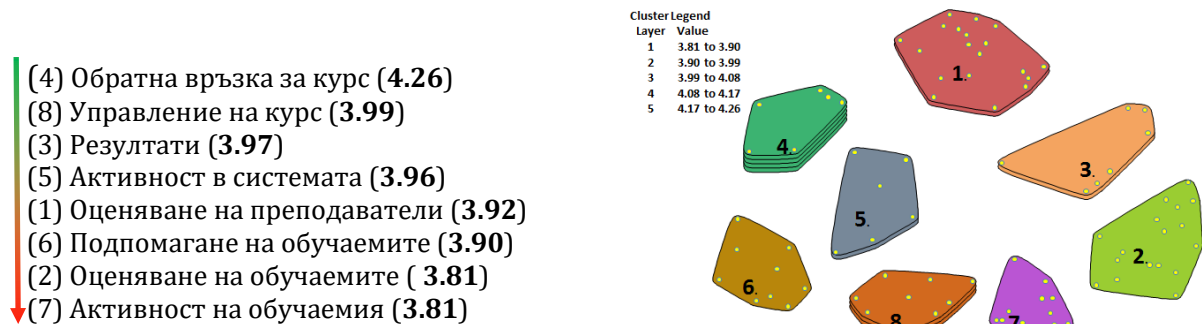
Показателен параметър за това, колко типична е всяка идея за групата, в която е разпределена, е нейната гранична стойност (bridging value). Това е параметър, чиято стойност се измерва в интервала $[0,1]$, като по-ниските стойности означават, че идеята е показателна за кълъстера, докато по-високите стойности показват, че мястото на идеята е на „границата“ на кълъстера, т.е. при по-голям брой кълъстери тя най-вероятно ще бъде част от друга група. Сред предложените идеи от експертите има представител с нулева гранична стойност и представител с максимална гранична стойност.

Освен за отделните идеи, средна гранична стойност се изчислява и за 8-те категории, в които те са групирани. Колкото по-малка е граничната стойност на кълъстера, толкова по-единодушни са експертите, че идеите в него трябва да бъдат заедно. И обратно, колкото по-висока е граничната стойност на даден кълъстер, толкова „по-обобщаващ“ е той по съответния критерий.

4.1.3. Оценка на идеите по критерий полезност/значимост

Диапазонът на средните оценки по критерий полезност/значимост варира от 3,10 до 4,60, като най-слаба оценка са получили две предложения: (30) *Справка каква част от курса е изминала и каква част остава в проценти* (M (средна оценка) = 3,10; SD (стандартно отклонение) = 1,3) и (35) *Да се показва статистика за активността на преподавател във форумите* (M = 3,10; SD = 0,8), а като най-полезно е оценено (39) *Графично представяне на конфликт на дати (за тест, домашно, ...) между курс при даден преподавател и други курсове на същите студенти* (M = 4,6; SD = 0,7).

От оценките на отделните идеи се изчислява и средна оценка за всяка категория от справки. От Фигура 3 се вижда, че като най-полезна е оценена категорията „Обратна връзка за курс“ с резултат 4,26 от 5, а на последно място по полезност са поставени групите „Оценяване на обучаемите“ и „Активност на обучаемия“.



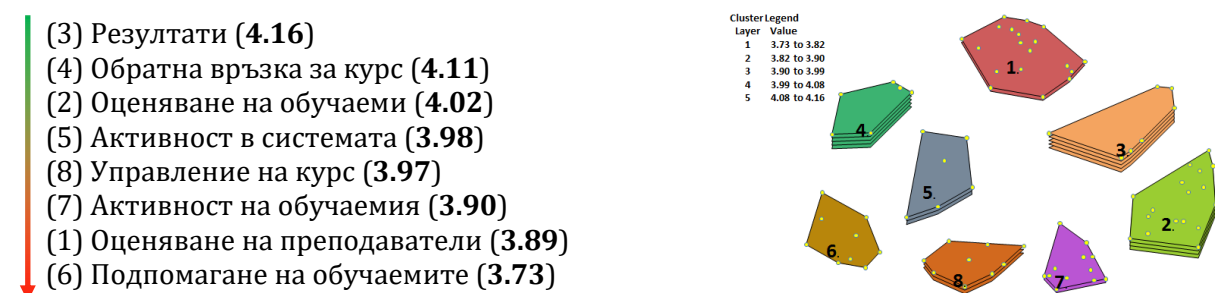
Фигура 3: Градиране на категориите по критерий полезност/значимост

Резултатите от оценките за полезност/значимост на по-късен етап са използвани за избор и приоритизиране на справките, които планираната система за учебни аналитики следва да предложи.

4.1.4. Оценка на идеите по критерий приложимост/осъществимост

Средните оценките на идеите, дадени от експертите по втория критерий приложимост/осъществимост, варират в диапазона [2,95 – 4,45]. Като най-трудно приложима е оценена идеята (63) *Да се показва прогнозно време за обявяване на резултати от контролно/домашно/изпит* ($M = 2,95$; $SD = 1,23$), а като най-лесна за реализация е посочена (24) *За всяко задание/дейност да се извежда списък на всички студенти, предали заданието/извършили дейността* ($M = 4,45$; $SD = 1,05$).

От оценките на идеите са изчислени средни оценки за всяка категория. От Фигура 4 се вижда, че като най-приложими са оценени справките в категория „Резултати“ – оценка 4,16 от 5, а като най-трудни за реализация са определени предложенията в категория „Подпомагане на обучаемите“ – оценка 3,73 от 5. По отношение на категориите разликата между минималната и максималната средна оценка не е голяма.

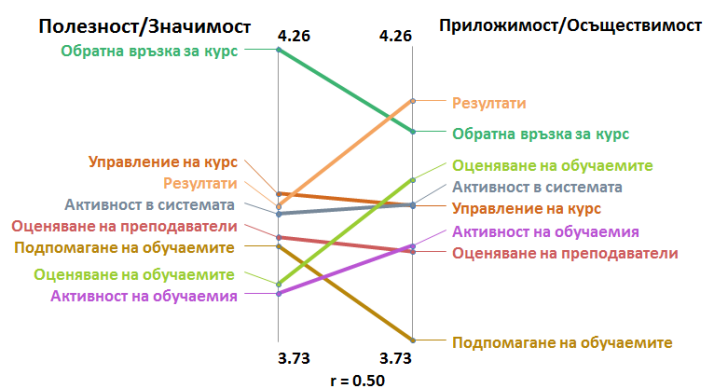


Фигура 4: Градиране на категориите по критерий приложимост/осъществимост

Резултатите от оценките за приложимост/осъществимост на по-късен етап са взети под внимание при избора и приоритизирането на справките, които планираната система за учебни аналитики следва да предложи.

4.1.5. Съпоставяне на оценките по двата критерия

Интерес представлява съпоставянето на средните оценки на предложените справки по категории по двата критерия. Някои от категориите са получили почти еднакво средно оценяване, например „Активност в системата“ (полезност/значимост: 3,96 от 5 – приложимост/осъществимост: 3,98 от 5) или „Управление на курс“ (полезност/значимост: 3,99 от 5 –приложимост/осъществимост: 3,97 от 5), докато справките в други категории са оценени като значително по-лесно приложими, отколкото полезни, например „Резултати“ (полезност/значимост: 3,97 от 5 – приложимост/осъществимост: 4,16 от 5) или „Оценяване на обучаемите“ (полезност/значимост: 3,83 от 5 – приложимост/осъществимост: 4,02 от 5) или по-полезни, отколкото лесно приложими, например „Подпомагане на обучаемите“ (полезност/значимост: 3,90 от 5 – приложимост/осъществимост: 3,73 от 5). Оценките за полезност са по-високи от оценките за приложимост (Фигура 5).



Фигура 5: Сравнение на средните оценки на категориите по двата критерия

4.1.6. Резултати

Оценките на предложените справки могат да бъдат обобщени в точкови диаграми, разделящи областта на 4 зони, спрямо средните стойности на оценките по двата критерия: полезност и приложимост (go-zone diagrams). В зелената зона или I квадрант се визуализират справките, чиито оценки са над средната стойност по двата критерия, в оранжевата зона или II квадрант се позиционират справките, които по критерий приложимост са получили над средната стойност, но по критерий полезност са оценени под средната стойност, в сивата зона или III квадрант са справките, които и по двата критерия имат оценка под средната за съответния критерий и в жълтата зона или IV квадрант са показани справките, които са оценени над средното ниво по критерий полезност, но под средното ниво по приложимост. Освен, че визуализират разпределението на справките във всяка категория като цяло, тези диаграми могат и да филтрират резултатите по роли: мениджър, преподавател и студент. Така от една страна да се виждат справките, които всяка роля е приоритизирала, а от друга - сеченията, т.е. справките, които би трябвало да се разработят за потребителите на СУО във всички роли.

В първата версия на системата за визуализиране на аналитики ще бъдат реализирани справките, които са оценени от експертите над средните стойности и по двата критерия. Справките, които са получили над средната оценка само по един

от двата критерия, ще бъдат ревизирани и реализирани на втора итерация. Справките, които са оценени под средната оценка и по двата критерия, ще бъдат преразгледани, преработени и предложени за ново оценяване. В Таблица 2 е показан броя на справките по категории и зони.

Таблица 2: Брой справки по категории и роли

Категория	Мениджър				Преподавател				Студент			
	I	II	III	IV	I	II	III	IV	I	II	III	IV
„Оценяване на преподавател“	9	2	7	1	11	3	5	1	6	3	5	5
“Оценяване на обучаемите“	9	4	2	2	4	5	3	5	5	5	6	1
„Резултати“	4	2	0	1	3	0	4	0	3	1	2	1
„Обратна връзка за курс“	3	1	2	0	3	2	1	0	1	2	1	2
„Активност в системата“	2	1	3	0	1	2	0	3	1	2	1	2
„Подпомагане на обучаемите“	3	1	3	2	2	3	3	1	4	2	1	2
“Активност на обучаемия”	4	2	3	3	5	1	4	2	5	1	4	2
“Управление на курс”	3	2	3	1	4	2	2	1	4	0	3	2

В тази глава са извлечени основните изисквания към интегрирана система за УА в два аспекта: (1) списък на всички справки, които заинтересованите страни биха искали да намерят в частта за УА, (2) за всеки тип от представени данни в тези справки е избрана визуализацията, определена като най-интуитивна, лесна за разчитане и бърза за разбиране.

4.2. Експертната гледна точка за избор на визуализации на учебните аналитики

В тази секция е описано проучване (Miteva & Stefanova, 2020b) на експертното мнение за типовете визуализации, с които синтезираните справки за СУО, описани в предходната секция, е най-удачно да бъдат онагледени.

4.2.1. Методология на проучването

В проучването участват 48 експерти с различна квалификация, опит и роля в СУО. Участниците попълват онлайн анкета „Визуализация на данни в СУО“, споделяйки своето професионално мнение за това кои визуализации онагледяват най-подходящо и точно данните в основните видове справки.

Списъкът на справки за раздела с учебни аналитики в СУО, дефиниран в 4.1 е анализиран допълнително по отношение на типовете данни на резултатите и целта на тяхната визуализация: композиция, сравнение, отношение или разпределение. В съответствие с класификацията на Abela са извлечени 12 основни типа данни за онагледяване. За всеки от тях е определен раздел в анкетата с кратко

описание какъв тип данни се очаква да бъдат визуализирани, на кого и за какво ще послужат. От 2 до 6 примерни визуализации са предложени за оценяване от 1 (неподходящи) до 5 (подходящи). Всеки раздел завършва с опция за коментари и идеи.

Събраните данни са обработени, анализирани и категоризирани според системните роли на респондентите, за да бъдат предложени виртуални табла с визуализации, които да удовлетворят възможно най-голям брой потребители.

4.2.2. Анализ на данните

В тази секция за описани детайли за предложените визуализации в анкетата и получените оценки според ролята на респондента в СУО. Коментирани са и препоръките, отправени към визуализациите.

Въз основа на анализа на данните от анкетата са предложени визуализации, които да онагледят резултатите от справките в частта за учебни аналитики според типа на данните. Тези визуализации са описани в следващия раздел.

4.2.3. Резултати

От целевата система за визуализация на учебни аналитики се очаква да предложи целесъобразно виртуално табло за потребителите с различна роля в СУО: мениджър, преподавател или обучаем. Изборът на визуализациите в него е направен на базата на две основни правила:

- Ако най-високата оценка за визуализация, която са дали експертите, използващи СУО в дадена роля, е над средната за диапазона, тази справка ще бъде включена в тяхното виртуално табло с избраната диаграма.
- Ако най-високата оценка, дадена от респондентите за дадена визуализация е под средната за диапазона, тази визуализация ще бъде пропусната в първата версия на системата и ще бъдат направени допълнителни изследвания.

4.2.4. Изводи

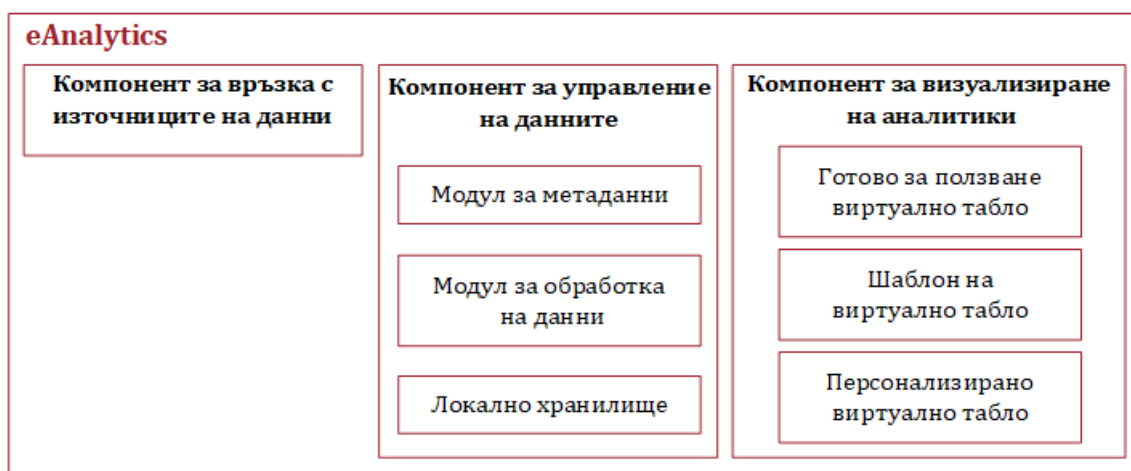
Сред предложените визуализации за експертна оценка има справки и отчети на статистически данни, представени в някои налични СУО и липсващи в други. Коментарите в анкетата показват, че отношението към познатите визуализации е доста крайно – една група е свикнала с тях, приема ги и ги препоръчва, а друга група е недоволна и иска същите резултати да бъдат представен с нов, по-интуитивен метод. От друга страна, дори и да не одобряват добре познати диаграми, респондентите трудно могат да си представят как точно ще работят новите визуализации в реално време и с реални данни. Междувременно по-опитните експерти предлагат практически идеи, които обогатяват резултатите от проучването. Може да се заключи, че новите видове визуализации трябва да бъдат представени по-атрактивно в работна среда и след това проучването да бъде повторено. Тази идея е заложена в плана за бъдещо развитие на системата. За по-добра визуализация и интерпретация на данните са приложени съвременни методи и техники за представяне на големи данни (Bikakis, 2018), като редуциране на данни, йерархична визуализация, прогресивно показване на резултатите, инкрементална и адаптивна обработка, кеширане и предварително извличане. В бъдещ момент могат да бъдат добавени 3D визуализации или дори виртуална реалност.

Следващата стъпка е създаване на модел на система за учебни аналитики, която чрез съвременни методи за визуализация да подпомага заинтересованите страни при вземането на правилни решения в точния момент.

Глава 5: Модел на система за визуализиране на учебни аналитики eAnalytics

В тази глава е описан архитектурен модел на система за визуализиране на УА, наречена eAnalytics. Целта на модела е да покрива максимално изискванията на потребителите, използващи СУО в трите основни роли: студент, преподавател или мениджър на учебно заведение.

На Фигура 6 са представени елементите на модела.



Фигура 6: Архитектура на модела

5.1. Потребители и роли

eAnalytics идентифицира няколко системни роли по подразбиране, на които се предоставят различни права върху наличните функции:

- **Мениджър.** Предоставя се на управленския екип на институцията да извлича аналитики относно участието както на обучаеми, така и на обучителите в даден курс, цялостното представяне, както и мнението и оценката на студентите за техните преподаватели.
- **Преподавател.** Дава възможност на преподавателите да следят цялостното провеждане на курс, да наблюдават индивидуалното представяне на всеки обучаем, да създават комплексен анализ на представянето на студентите.
- **Студент.** Роля за обучаемите, с която да следят своите резултати, успехи и слаби места и да получават обратна връзка как да подобрят собственото си представяне.
- **Администратор.** Той е отговорен за инсталирането, конфигурирането и управлението на системата, има пълни права на достъп.

Позволява се дефиниране на нови роли и делегиране на съответни права.

5.2. Компоненти на системата

eAnalytics е самостоятелна веб-базирана система с архитектура на хъб между свързани учебни системи и съхранява данни за процесите на преподаване, учене и управление на обучението, за да даде цялостен поглед върху напредъка на обучаемите. Събирането на повече данни за по-прецизна оценка се реализира чрез обединяване на данни от множество източници на учебни аналитики. С помощта на API системата се свързва с различни СУО и бази от данни и чрез метаданни описва какви аналитики ще предоставят.

Системата се състои от три основни компоненти: (1) за връзка с източниците на данни, (2) за управление на данните и (3) за визуализиране на аналитики.

5.2.1. Компонент за връзка с източниците на данни

Компонентът за връзка с източниците на данни отговаря за осигуряване на връзка и достъп до източниците съобразно конкретния потребител, ролята и правата му на достъп в различните СУО. Метаданните за този достъп се съхраняват в системата и се използват за извличане на данни от СУО

5.2.2. Компонент за управление на данните

Този компонент включва три модула: (1) модул за метаданни, (2) модул за обработка на данните и (3) локално хранилище.

Модул за метаданни

След като бъдат извлечени метаданни и описани съответните връзки, този модул изгражда йерархичен модел на метаданните, структуриран по роля в системата, достъп до курс и предоставени данни. Основната единица, в която се групират метаданните в eAnalytics е курс, като достъпът до аналитиките зависи от ролята на потребителя в системата.

Модул за обработка на данни

За някои от аналитиките, предимно свързани с комбинирани данни от различни източници, е необходима допълнителна обработка в самата система. eAnalytics предлага набор от стандартни формули, математически и статистически функции за обработка и трансформиране на данни, за анализ на текст, за параметризиране на данните.

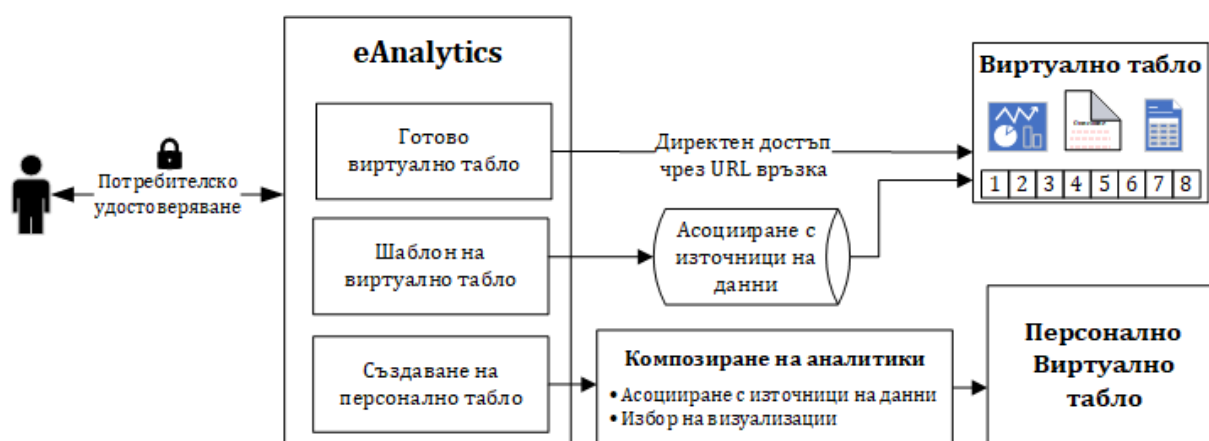
Локално хранилище

Извадки от големите обеми от данни могат да бъдат импортирани и съхранявани в локално хранилище. Такива могат да бъдат използвани, когато не е необходимо предаването на данни от източника към системата да се реализира в реално време. Тогава могат да се експортират и импортират файлове с формат .csv или .xlsx.

5.2.3. Компонент за визуализиране на аналитики

Компонентът предоставя избор между няколко метода за визуализация на аналитики: (1) чрез готово виртуално табло, (2) чрез шаблон на виртуално табло или (3) чрез композиране на персонални аналитики. При първия метод виртуалното табло съдържа информация за източниците на данни и може да бъде

споделено само за разглеждане или с допълнителна възможност за редакция. Втората опция предоставя възможност потребителят сам да избере/зададе източниците на данни, за да получи съответните визуализации. Полученото виртуалното табло остава в профила на потребителя. Третата опция директно препраща към екран за композиране на персонализирани аналитики, в който най-напред трябва да бъдат избрани източниците на данни, след това самите данни и накрая визуалните обекти, които ще представят данните. На база на избраните типове данни ще бъдат препоръчани „подходящи“ визуализации, но изборът остава в ръцете на потребителя (Фигура 7).



Фигура 7: Компонент за визуализация на аналитики

Визуализация чрез готово виртуално табло

Виртуалното табло, до което потребителят получава достъп, е напълно готово за разглеждане. Основното му предимство е, че не изисква допълнителни настройки, нито познаване на структурата на данните и връзките между тях. Съдържанието му варира в зависимост от системната роля на потребителя. Ако потребителят има присвоени повече от една роли, той може да избира с коя от тях ще участва в дадения момент.

Таблото се състои от осем секции, по една за всяка от категориите, описани в глава 4: (1) Оценяване на преподаватели, (2) Активност на обучаемия, (3) Оценяване на обучаемите, (4) Управление на курс, (5) Резултати, (6) Подпомагане на обучаемите, (7) Активност в системата и (8) Обратна връзка за курс. Списъкът с аналитични справки и визуализации във всяка категория е различен и съответства на текущата роля и експертното мнение за избор на визуализации, описано в 4.2.

Компонентът за визуализиране на аналитики позволява превключване между разработените по подразбиране визуализации за различните роли, както и отразяване на настройките в профила на потребителя. Допълнително са предоставени опции за филтриране на данните и експортиране на резултатите в табличен и/или графичен формат.

Визуализация чрез шаблон на виртуално табло

Шаблонът е структуриран като готовото виртуално табло, съдържа всички аналитики и подходящите визуализации за тях (Глава 4), пази в себе си модела на метаданните и връзките между тях, но липсва свързаност с конкретни източници на данни и локалното хранилище е празно. За да могат да се визуализират аналитиките, най-напред шаблонът трябва да се асоциира с източници на данни. След това *компонентът за връзка с източниците на данни и модулът за метаданни* допълват и актуализират метаданните. Следва извличане, обработка на данните и визуализиране на съответните аналитики във виртуалното табло.

Този метод за визуализиране има редица предимства. Виртуалното табло е универсално, може да работи с различни източници, които поддържат еднотипен модел на данните. За eAnalytics това означава, че разработено за една конкретна институционална структура, то може да бъде приложено за множество аналогични структури.

Визуализиране чрез създаване на персонализирано табло

Създаването на персонално виртуално табло дава възможност на потребителя сам да дефинира какви аналитики иска да включи и какви визуализации да използва за тях. Персоналното виртуално табло се създава в следната последователност от действия в системата: (1) избиране на данни от описаните в *модул метаданни*, (2) избиране на визуализации, (3) персонализиране на параметрите.

Данните, които се предоставят за композиране на аналитики, са описани ясно и интуитивно в системата и дават точна представа каква информация съдържат. Могат да бъдат достъпни от различни перспективи: данни за студенти, данни за преподаватели, данни за курсове.

Всеки потребител може да персонализира общата визия на системата за своя профил. Роля администратор определя глобалната визия на цялата система.

5.3. Мобилност, сигурност и етични норми

Без значение дали потребителите се свързват от настолен компютър, лаптоп, таблет или смартфон, отчетите за анализ се предлагат/адаптират във възможно най-добрия формат за съответното устройство. Системата е достъпна чрез приложения за Android и iOS.

С цел да се намали риска от уязвимост при съхранението на чувствителни данни моделът предвижда обръщение към съответните СУО само при заявка и за извличане само на поисканите данни, предоставени срещу съответен акаунт. При необходимост от публикуване на резултати пред трети лица, данните се анонимизират.

Моделът предвижда прилагане на етичен кодекс, съобразен с изискванията на организацията, която ще ползва съответната система за учебни аналитики и достъпен в явен вид през системата.

Глава 6: Създаване на прототип

В тази глава е представена работата по шестата задача – създаването на прототип на система за визуализиране на УА. Демонстрирано е приложение на модела за системна роля мениджър с готово виртуално табло за категория „Оценяване на преподавател“. В прототипа е реализиран и шаблон на виртуално табло, който може да бъде използван многократно, като потребителят сам избира с кои източници на данни да го свърже. Шаблонът пази информация за модела на данните, за създадените визуализации в него, но не съдържа конкретни данни.

6.1. Избор на средство за реализация на виртуално табло

Съществуват редица съвременни технологии и инструменти за визуализиране на големи обеми от данни, които навлизат все по-широко в сферата на образованието. Някои от тези инструменти предлагат лесен и удобен интерфейс за потребителя, други изискват сериозни технологични и програмни познания. Част от тези инструменти са тясно специализирани за конкретни типове визуализация, например географски карти, други предоставят разнообразие от стилове и са отворени за добавяне на нови.

На базата на описаните предимства и недостатъци на инструментите за създаване на виртуално табло в глава 1 е направен избор за създаване на прототипа да бъде използвана платформата Microsoft Power BI, като покриваща в най-голяма степен на описаните критерии за избор на платформа.

6.2. Данни и метрики, използвани в прототипа

Прототипът, реализиращ модела, е демонстриран върху реален проблем, който съществува с обработката и визуализирането на големи обеми от данни, свързани с обратната връзка от студентите в курсовете за обучение в Софийски университет „Св. Климент Охридски“.

В края на всеки учебен семестър в СУ „Св. Климент Охридски“ се провежда „анкета за оценка на курс“, чрез която студентите дават обратна връзка за учебните дисциплини, които са изучавали, каква е била посещаемостта на лекции и упражнения, работата на своите преподаватели и асистенти, споделят мнения и препоръки. Тази анкета съдържа около 30 въпроса, разделени в 4 секции. В първата секция студентите оценяват по скала от 1 до 6 колко интересен и полезен е бил за тях учебният материал, доколко е допринесъл за тяхната подготовка и професионално развитие. Във втората секция всеки студент споделя каква част от лекциите и упражненията (ако е приложимо) е посетил. В третата и четвъртата секция в рамките на 11 въпроса от тип скала с оценка от 1 до 6 или да/не/не мога да преценя, студентите оценяват работата на преподавателя и асистента – дали поднасят достъпно учебния материал, дали създават творческа обстановка и стимулират самостоятелно мислене, провеждат ли редовно занятия, дали одобряват методите на преподаване и системата за контрол и оценка, дали студентите биха препоръчали учебната дисциплина на свои колеги. Всеки от

въпросите в част III и IV се оценява с определено тегло и по регламентирана формула в *Наръчник по качество на СУ* формира част от атестационната оценка на преподавателя. До момента резултатите от тази анкета се обработват ръчно и това затруднява пълноценния им анализ и извличането на информация.

Виртуалното табло, описано в тази глава, е разработено по модел на данните от тези анкети.

6.3. Импортиране на данните и свързването им в модела

За целите на демонстрацията данните се получават като „Работна книга на Ексел“ и „Текст/CSV“, като предварително за оформени в 3 различни групи, симулиращи 3 различни източника на данни.

Изграден е модел на метаданните, в който са указани връзките между отделните набори от данни. Възможностите за релации са (1) един към един, (2) един към много, (3) много към един и (4) много към много. Средата предоставя възможност отношенията да се дефинират визуално чрез функцията влачене и пускане.

6.4. Създаване на виртуално табло и избор на визуализации

Съгласно модела на eAnalytics виртуалното табло за оценка на преподавател, достъпно за роля „мениджър“, предоставя на потребителя 8 справки. Визуализираните справки са групирани в четири секции/страници, като в началото е добавена една събирателна, с въвеждащи обяснения какво потребителят да очаква да намери в това виртуално табло и какви са критериите за оценка.

6.5. Създаване на шаблон на виртуалното табло

За да може създаденото виртуално табло да бъде използвано многократно, но с различни източници на данни, които поддържат същия формат и модел на данните, системата предлага възможността за създаване на шаблон на виртуалното табло. Ако в него източникът на данни се зададе като параметър, то при следващото отваряне, потребителят ще може да избира с какви данни иска да използва шаблона.

В прототипа източниците на данни са зададени като параметри. При стартиране се задават стойности на параметрите и всички данни ще получат реален път до съответния източник

6.6. Създаване на приложение и споделяне в онлайн магазин

За целите на прототипа е създадено приложението eAnalytics (eAnalytics_app), което е публикувано в Power BI apps.

Глава 7: Тестване на прототипа

В тази глава е описано експерименталното тестване на създадения прототип на eAnalytics, в отговор на седмата поставена задача. Проведени са функционални

тестове за съответствие на реализираните със заложените изисквания към прототипа и потребителски тестове чрез прилагане на протокола “Мисля на глас” (Think aloud protocol) (Lewis, 1982).

7.1. Предварителна подготовка на данните

Преди да започне съществената обработка, данните са почистени от некоректни или частично попълнени записи. Справките, генерирани с данни от СУО съдържат чувствителни данни и за целите на демонстрацията е приложена деидентификация за прикриване на личните данни и минимизиране на риска от непреднамерено разкриване на преподаватели и информацията за тях. Използван е инструментът с отворен код Amnesia⁷, който следва указанията на GDPR за псевдо-анонимизация. Чрез него връзката с оригиналния набор от данни е прекъснатата. На псевдо-анонимизация са подложени идентификаторите *име, презиме, фамилия* и *учебна дисциплина*, които в комбинация или поотделно биха помогнали да се разпознае оценявания преподавател.

7.2. Функционално тестване на прототипа

В тази секция са описани проведените тестове на разработените функционалности на eAnalytics: (1) отваряне на шаблон за виртуално табло и конфигурирането му за работа с реални източници на данни, (2) тестване на готово виртуално табло и (3) създаване на персонализирано виртуално табло.

7.2.1. Тестване на шаблон за виртуално табло

При отваряне на шаблона, най-напред се появява диалогов прозорец, подканващ потребителя да избере (или въведе) реални източници на данни, които да бъдат визуализирани във виртуалното табло. В тествания шаблон потребителят е улеснен директно да избира от списък с предварително описани СУО. Моделът прочита посочените адреси, свързва се с източниците и зарежда данните.

Най-напред се показва началният екран (Фигура 8), в който са описани анализирания период, критериите за оценка на студентското мнение и видовете предоставени справки с връзки към по-детайлното им представяне.



Фигура 8: Виртуално табло „Студентско мнение за преподавателите“

⁷ Amnesia Anonymization Tool, <https://amnesia.openaire.eu/index.html>, последен достъп: 01.02.2023

Следвайки връзките или навигацията между отделните страници, постепенно могат да бъдат разгледани детайлно всичките секции на виртуалното табло

7.2.2. Тестване на готово виртуално табло

Готовото виртуално табло, публикувано като приложение в онлайн магазин, е тествано и използвано от регистриран потребител на СУ в Power BI service, с когото то е споделено. Достъпът до него е установен (1) чрез предоставена URL връзка при споделяне, (2) чрез сканиране на QR код и (3) чрез търсене по име като „eAnalytics_app“ в онлайн магазина на Power BI apps.

Приложението е изтеглено в профила на потребителя, инсталирано е и е достъпно за онлайн ползване по всяко време (Фигура 9).



А) eAnalytics, достъпена от компютър



Б) Мобилна версия на eAnalytics

Фигура 9: eAnalytics като приложение

7.2.3. Тестване на създаване на персонализирано виртуално табло

eAnalytics е споделена с тестов потребител с права за редакция. Този потребител може да генерира свои аналитики и персонализирани визуализации с предоставените данни. Той е създал нова справка, в която (1) е маркирал да използва данните за брой преподаватели със средна положителна, задоволителна и незадоволителна атестация за всички анализирани семестри, (2) като визуален обект е избрал кръгова диаграма, (3) персонализирал е настройките на визуализацията и е онагледил съотношението на всички преподаватели според атестацията за целия анализирани период (Фигура 10).



Фигура 10: Тестване на създаване на персонализирано табло

От направените тестове може да се заключи, че реализираните функционалности на прототипа, описани в глава 6, работят както са заложиени в прототипа.

7.3. Потребителски тестове с протокол „Мисля на глас“

В тестването на eAnalytics се включиха 10 експерти от Факултета по математика и информатика на СУ „Св. Климент Охридски“, които декларираха, че използват активно СУО в работата си повече от 10 години, като един от тях е включен в комисията по атестиране на ФМИ. С всеки от тях предварително беше споделено приложението от онлайн магазина на Power BI и всеки участник трябваше да изпълни следните задачи за тестване на готовото виртуално табло в уеб формат и мобилна версия:

1. Намиране и изтегляне на приложението eAnalytics в онлайн магазин.
2. Стартиране на приложението.
3. Обхождане на всички страници, разглеждане на предложените справки и ориентация в навигацията.
4. Превключване между различните видове визуализации, налични за една и съща справка.
5. Откриване на най-високата и най-ниската оценка на преподавател за всеки от показаните 5 семестъра. (Без да се уточнява в коя справка да се търси информацията.)
6. Извличане на три различни списъка на преподаватели: с незадоволителна, задоволителна и положителна атестация за произволно избран семестър. (Без да се уточнява в коя справка да се търси информацията.)

Всеки участник беше инструктиран да споделя на глас всяка мисъл, която му хрумва, докато изпълнява поставените задачи.

Резултати от наблюдението

С първата и втората задача всички се справиха изключително бързо. Навигацията беше оценена като лесна и интуитивна, а превключването между различните визуализации – бързо откриваемо и логично. Задачи 5 и 6 бяха изпълнени с лекота. Единствената забележка беше свързана отново с показването на бутон за изтегляне в явен вид, а не само когато обектът е на фокус, т.е. мишката е върху него.

Първите впечатления бяха изразени с положителна превъзбуда и възклицания от типа: „О, идеално, това ще бъде много полезно за шефовете!“, „Охо, интерфейсът е адаптиран за потребители със специални потребности!“, „Това с цветовете е много добро, няма нужда да гледаш числата, за да се ориентираш!“, „О, и това е много полезно!“ и др. След първоначалната еуфория беше направена по-критична оценка и списък с препоръки.

След уеб приложението, за тест беше предложена мобилната версия. Тя беше оценена като по-интуитивна, по-удобна за работа и с по-добър дизайн.

Тестването на прототипа показва експериментално, че обемните данни, статистическите изчисления и тяхното графично визуализиране ускоряват

многократно възприемането им и дават възможност на участниците в обучението да отговорят поне на част от поставените въпроси в началото.

Заключение

Дисертацията „Визуализация на големи обеми от данни“ изследва възможностите за повишаване на ефективността на процесите на учене, преподаване и управление на обучението чрез използване на съвременни методи за визуализиране на учебни аналитики. Нейната цел е проектирането на система, която чрез подходящи визуализации да подпомага обучаеми, преподаватели и мениджъри на учебни заведения за вземане на правилни решения в точния момент.

Постигнати резултати

Приносителите на дисертацията са както следва:

Проучено е използването на учебни аналитики в образованието (Глава 1). Анализирани са метриците и ключови показатели за ефективност, използвани в СУО. Проучени са съвременните методи за събиране на данни от системите за управление на обучението, описващи успеха и прогреса на обучаемите, тяхната активност и отговорност към поставените задачи, отношението им към обучението и социалната активност. Разгледани са по-известните модели за прилагане на учебни аналитики в обучението. Коментирани са етичните правила за събиране и обработка на лични данни, регламентирани от правния инструмент за защита GDPR.

Анализирани са съвременните методите за визуализация на различни категории от данни: статични, интерактивни и големи. Дискутирани са средства за визуализация на учебни аналитики, използвани активно в MOOC курсове.

Изследвани са платформите за създаване на виртуално табло. Разгледани са водещите платформи за анализ и бизнес интелигентност според класацията на Gartner за 2022.

Проведени са експерименти с прилагане на различни УА методи и подходи за оптимизиране на учебния процес (Глава 2) с цел (1) подобряване на качеството на обучение за студентите, (2) подобряване на качеството на преподаване от страна на преподавателите и (3) подобряване на цялостната организация на обучението на ниво обучаваща организация.

Проведени са експерименти с прилагане на различни методи за визуализиране на учебни аналитики (Глава 3), като Moodle, GISMO, Canvas analytics, LARAE, MICI (последните две са създадени по проект weSPOT). Вследствие е идентифициран набор от предизвикателства, на които трябва да се намерят решения при създаването на една съвременна система за визуализация.

Дефинирани са основните изисквания към интегрирана система за учебни аналитики в два аспекта: (1) от една страна е изведен списък на всички справки, които заинтересованите страни (обучаеми, преподаватели и мениджъри на учебни

заведения) биха искали да намерят в частта за учебни аналитики, (2) от друга страна за всеки тип представени данни в тези справки е избрана визуализацията, определена като най-интуитивна, лесна за разчитане и бърза за разбиране (Глава 4). Списъкът със справки е обособен в осем категории, като избраните отчети и журнали в него варират за различните роли потребители на СУО, както и визуализациите за всеки от тях.

На базата на дефинираните изисквания е създаден модел на система за визуализиране на учебни аналитики, наречен eAnalytics. Той предвижда събиране и комбиниране на данни за обучаемите от повече от една СУО, което да позволи по-пълнен анализ и по-точни резултати за статуса на отделния обучаем и преподавател. Визуализирани подходящо, тези резултати подпомагат вземането на правилни решения и ответни мерки за индивидуално подпомагане или насърчаване за по-добри постижения (Глава 4).

На базата на модела е създаден прототип, който реализира виртуално табло, визуализиращо студентската оценка за преподавател, предназначено за потребител с роля мениджър. Създаден е шаблон на таблото. Паралелно е създадено приложение, споделено през онлайн магазина на Power BI apps.

Накрая са извършени успешни тестове на шаблона и на онлайн приложението.

Изпълнени са всички поставени задачи в дисертацията, а оттам е постигната и целта да бъде създадена система, която чрез съвременни визуализации за големи обеми от данни да подпомогне обучаеми, преподаватели и мениджъри за повишаване на качеството на обучението.

Насоки за бъдещо развитие

Описаната в тази дисертация система за визуализиране на учебни аналитики ще продължи да бъде развивана в няколко посоки.

Като начало тя се базира на проучване какви справки за визуализиране на учебни аналитики експертите считат за полезни и необходими. За някои от тези справки системите за обучение събират и съхраняват достатъчно данни, за да бъдат реализирани. За други отчети е необходимо допълнително планиране от една страна какви данни да бъдат събирани, от друга страна как, с какви средства. За някои от предложените идеи, например свързани със спазване на предварителен график и уведомявания за предстоящи събития, е необходимо такъв примерен график да бъде изискан от преподавателите в началото на курса и да бъде за заложен в съответната СУО.

Предстои да бъде проведено още едно проучване за използване на съвременни методи за визуализация на данни, както е отбелязано в глава 4.. При него най-новите методи за визуализация на големи обеми от данни, приложени към УА, ще бъдат представени в реална среда, за да могат респондентите да оценят по-меродавно предимствата и предизвикателствата пред всеки от тях. За целите на това проучване ще може да бъде използвана текущата версия на eAnalytics.

Предвид чувствителността на данните в СУО, които се обработват и визуализират в учебните аналитики, една от насоките за доразвиване на eAnalytics е добавяне на локален за университета сървър, на който да бъдат хоствани виртуалните табла. Благодарение на него данните, които обработва и визуализира eAnalytics, няма да напускат мрежата на Софийски университет. Тяхното споделяне ще бъде с потребители от самия университет или в случай на външни потребители – чрез протоколи, гарантиращи сигурност и защита на данните. Към сървъра да бъде добавен „шлюз“ (gateway) приложение, което да позволи данните да останат на територията на университета дори тогава, когато виртуално табло бъде публикувано в облак. В допълнение ще бъде добавена и услуга за вграждане на виртуалните табла на eAnalytics в самите СУО.

Съвременните технологии за обработка на големи данни и визуализация на изображения непрекъснато се развиват. Изискванията към СУО непрекъснато се увеличават, разработват се нови функционалности, натрупват все повече и повече данни. Очакванията към визуализациите на учебни аналитики също ще растат. За да бъде полезна, eAnalytics ще бъде актуализирана периодично, за да отговаря на изискванията на своите потребители: преподаватели, студенти и мениджъри.

Авторска справка

Приноси на дисертационния труд

Приносите на дисертацията могат да бъдат резюмират като:

Научноизследователски приноси

1. Изследвани са учебните аналитики в образованието, използваните метрики и ключови показатели за ефективност, методите за събиране на данни за обучаеми и моделите за прилагане на учебни аналитики и тяхната роля за повишаване на качеството на преподаване и учене. Коментирани са етичните норми и защитата на личните данни при прилагането на учебни аналитики (Глава 1).
2. Проучени са методите за визуализация на големи обеми от данни и средствата за визуализиране на учебни аналитики (Глава 1, Публикации 1, Публикация 2).
3. Анализирани са съвременните платформи за създаване на виртуално табло (Глава 1).
4. Направен е експериментален анализ на прилагането на различни методи и подходи за УА с цел оптимизиране на учебния процес и са извлечени предизвикателствата за създаване на система за визуализиране на учебни аналитики (Глава 2, Публикация 3).
5. Изследвани са средства за визуализиране на учебни аналитики (Глава 3, Публикация 4).

6. Анализирани и систематизирани са учебните аналитики, които потребителите на СУО очакват да намерят в секцията за учебни аналитики (Глава 4, Публикация 5).
7. Анализирани и систематизирани са визуализациите на учебни аналитики в СУО, които потребителите е различни роли (обучаеми, преподаватели и мениджъри) намират за най-интуитивни и лесни за разчитане (Глава 4, Публикация 6).

Научноприложни приноси

1. Разработен е модел на система за визуализиране на учебни аналитики, който да покрива максимално изискванията на потребителите на СУО (Глава 5).
2. Създаден е прототип на модела на система за визуализиране на учебни аналитики. Демонстрирано е приложение на модела за системна роля мениджър в категория „Оценяване на преподавател“ (Глава 6).
3. Тестван е експериментално създаденият прототип на виртуално табло, част от системата за учебни аналитики (Глава 7).

Публикации

В дисертационния труд са използвани части от публикациите на автора:

Към глава 1:

1. Miteva, D., Stefanova, E. (2018). From Big Data to Big Decisions via the Bridge of Learning Analytics Visualization, Eight International Scientific Conference 'COMPUTER SCIENCE' 2018, pages:85–91, ISBN:978-619-167-359-9
2. Митева, Д., Стефанова, Е., Визуализиране на учебни аналитики за по-ефективно обучение, Седма национална конференция "Електронното обучение във висшите училища", издателство: Университетско издателство "Св. Климент Охридски", 2018, стр.:143-150, ISBN:978-954-07-4509-1

Към глава 2:

3. Miteva, D., Stefanova, E., Stefanov, K. (2017) e-Analytics for e-Learning, International Journal of Human Capital and Information Technology Professionals (IJHCITP), vol:8, issue:4, 2017, pages:1-13, ISSN (print):1947-3478, ISSN (online):1947-3486, <https://doi.org/10.4018/IJHCITP.2017100101>, SJR (0.28 – 2017)

Към глава 3:

4. Miteva, D., Stefanova, E. (2016), Driving teaching interventions with analytics, International Conference END 2016 Book of Proceedings, editor/s:Mafalda Carmo, Publisher:World Institute for Advanced Research and Science (WIARS), Portugal, 2016, pages:349-353, ISBN:978-989-99389-8-4
<http://end-educationconference.org/wp-content/uploads/2016/07/END-2016-Book-of-Proceedings.pdf>

Към глава 4:

5. Miteva, D., Stefanova, E. (2020). Design of Learning Analytics Tool: The Experts' Eyes View. Proceedings of the 12th International Conference on Computer Supported Education – Volume 2: CSEDU, 307–314.
<https://doi.org/10.5220/0009395503070314>, ISBN978-989-758-417-6
6. Miteva, D., Stefanova, E. (2020). Experts' perspective for selecting learning analytics visualizations. 2020 International Conference Automatics and Informatics (ICAI), 1–6. <https://doi.org/10.1109/ICAI50593.2020.9311341>

Научни проекти, свързани с дисертационния труд

1. Improving online and blended learning with educational data analytics (ILEDА), по програма Еразъм+, 2021-1-BG01-KA220-NED-000031121, 2021-2023
2. Използване на високотехнологични средства при изграждане на компетентностни модели в обучението, №80-10-18/18.03.2020 г., Фонд за научни изследвания на Софийски университет „Св. Климент Охридски“
3. Национална научна програма "Информационни и комуникационни технологии за единен цифров пазар в науката, образованието и сигурността" (ИКТ в НОС), с финансовата подкрепа на Министерството на образованието и науката на Република България, 2018 – 2021
4. Използване на високи технологии при прилагане на изследователски подход за обучение по природни науки в средното училище, №80–10-90/19.04.2018 г., Фонд за научни изследвания на Софийски университет „Св. Климент Охридски“
5. Изследователският подход в обучението в областта на високите технологии като приложение на съвременните информационни технологии, №80–10-217/24.04.2017, Фонд за научни изследвания на Софийски университет „Св. Климент Охридски“
6. Изследване на приложимостта в практиката на нови технологични средства и подходи за обучение, №30/31.03.2015 г., Фонд за научни изследвания на Софийски университет „Св. Климент Охридски“
7. Методи и технологии за подобряване на достъпа до електронни услуги – ASSETS (Automatic Service diScovery, sEmantic annoTation and compoSition), FNI02–68, 2015 – 2017
8. Working Environment with Social and Personal Open Tools for inquiry based learning (weSPOT), EC-FP7, FP7-ICT-2011-8-318499, 2012 – 2015

Декларация за оригиналност

Декларирам, че представената във връзка с провеждането на процедура за придобиване на образователната и научна степен „доктор“ в Софийски университет “Св. Климент Охридски“ дисертация на тема: „Визуализация на големи обеми от данни“ е мой труд.

Цитиранията на всички източници на информация, текст, илюстрации, таблици, изображения и други са обозначени според стандартите.

Резултатите и приносите на проведеното дисертационно изследване са оригинални и не са заимствани от изследвания и публикации, в които нямам участие.

Библиография

1. Abela, A. V. (2010). *The presentation: A story about communicating successfully with very few slides*. CreateSpace.
2. Bikakis, N. (2018). *Big Data Visualization Tools*.
3. Caldarola, E. G., & Rinaldi, A. M. (2017). Big Data Visualization Tools: A Survey. *Proceedings of the 6th International Conference on Data Science, Technology and Applications*, 296–305. <https://doi.org/10.5220/0006484102960305>
4. Chen, Q. (2018). *Visual analytics and storytelling of data from massive open online courses* [The Hong Kong University of Science and Technology]. <https://repository.hkust.edu.hk/ir/Record/1783.1-95778>
5. Few, S. (2006). *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media, Incorporated.
6. Florea, D., & Florea, S. (2020). Big Data and the Ethical Implications of Data Privacy in Higher Education Research. *Sustainability*, 12(20). <https://doi.org/10.3390/su12208744>
7. Hout, N. (2020, July 6). *10 Accurate Training Metrics To Always Include In Your Learning Analytics Report*. <https://elearningindustry.com/training-metrics-to-include-in-learning-analytics-report>
8. Iliinsky, N., & Steele, J. (2011). *Designing Data Visualizations: Representing Informational Relationships 1st Edition* (1st edition). O'Reilly Media; 1st edition. <https://www.amazon.com/Designing-Data-Visualizations-Informational-Relationships/dp/1449312284>
9. Kane, M. A., & Rosas, S. R. (2017). *Conversations About Group Concept Mapping: Applications, Examples, and Enhancements* (Paperback). SAGE Publications, Inc. <https://lead.to/amazon/com/?op=bt&la=en&cu=usd&key=1506329187>
10. Khalil, M., & Ebner, M. (2016, March). *What is Learning Analytics about? A Survey of Different Methods Used in 2013-2015*.
11. Khan, M., & Khan, S. S. (2011). Data and Information Visualization Methods, and Interactive Mechanisms: A Survey. *International Journal of Computer Applications*, 34(1), 1–14.
12. Lewis, C. (1982). *Using the “thinking-aloud” method in cognitive interface design*. IBM Thomas J. Watson Research Division. <https://books.google.bg/books?id=F5AKHQAACAAJ>
13. Mediratta, A. (2015). Big Data: Terms, definitions, and applications. *EMC Proven Professional Knowledge Sharing*.
14. Miteva, D., & Stefanova, E. (2020a). Design of Learning Analytics Tool: The Experts' Eyes View. *Proceedings of the 12th International Conference on Computer Supported Education - Volume 2: CSEdu*, 307–314. <https://doi.org/10.5220/0009395503070314>

15. Miteva, D., & Stefanova, E. (2020b). Experts' perspective for selecting learning analytics visualizations. *2020 International Conference Automatics and Informatics (ICAI)*, 1–6. <https://doi.org/10.1109/ICAI50593.2020.9311341>
16. Pauwels, K., Ambler, T., Clark, B. H., LaPointe, P., Reibstein, D., Skiera, B., Wierenga, B., & Wiesel, T. (2009). Dashboards as a Service: Why, What, How, and What Research Is Needed? *Journal of Service Research*, *12*(2), 175–189. <https://doi.org/10.1177/1094670509344213>
17. Rouse, W. B., Pennock, M. J., Oghbaie, M., & Liu, C. (2017). Interactive visualizations for decision support: Application of Rasmussen's abstraction-aggregation hierarchy. *Applied Ergonomics*, *59*(Pt B), 541–553. <https://doi.org/10.1016/j.apergo.2016.03.006>
18. Salihoun, M. (2020). State of Art of Data Mining and Learning Analytics Tools in Higher Education. *International Journal of Emerging Technology in Learning*, *15*(21). <https://www.learntechlib.org/p/218354/>
19. Schwendimann, B. A., Rodríguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2017). Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies*, *10*(1), 30–41. <https://doi.org/10.1109/TLT.2016.2599522>
20. Shi, C., Fu, S., Chen, Q., & Qu, H. (2014). VisMOOC: Visualizing video clickstream data from massive open online courses. *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 277–278. <https://doi.org/10.1109/VAST.2014.7042528>
21. Siemens, G. (2010). What are Learning Analytics? [Internet Archive]. *ELEARNSPACE*. <https://web.archive.org/web/20180630115729/http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics>
22. Siemens, G., Gašević, D., Haythornthwaite, C., Dawson, S., Buckingham Shum, S., Ferguson, R., Duval, E., Verbert, K., & Baker, R. S. (2011). *Open Learning Analytics: An integrated modularized platform* [WorkingPaper].
23. Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A systematic literature review and research agenda. *Computers & Education*, *122*, 119–135. <https://doi.org/10.1016/j.compedu.2018.03.018>
24. Wang, L., Wang, G., & Alexander, C. A. (2015). Big Data and Visualization: Methods, Challenges and Technology Progress. *Digital Technologies*, *1*(1), 33–38. <https://doi.org/10.12691/dt-1-1-7>
25. Wang, R. (2012). *Monday's Musings: Beyond The Three V's of Big Data—Viscosity and Virality*.
26. Whitehouse, A. J. O., Maybery, M. T., & Durkin, K. (2006). The development of the picture-superiority effect. *British Journal of Developmental Psychology*, *24*(4), 767–773. <https://doi.org/10.1348/026151005X74153>
27. Митева, Д., & Стефанова, Е. (2018). Визуализиране на учебни аналитики за по-ефективно обучение. 143–150.