



Софийски университет „Св. Кл. Охридски”

Факултет по математика и информатика

Катедра „Софтуерни технологии”



ДИПЛОМНА РАБОТА

на тема

**„Система за синтез и визуализация на
текст с разнообразен произход
(Ultimate Speaker) „**

Дипломант: Асен Георгиев Асенов

Специалност: Разпределени системи и мобилни технологии

Факултетен номер: M24072

**Научен ръководител:
доц. д-р Милен Петров**

София, 2015г.

Съдържание

Глава 1. Увод (3-5стр.).....	4
1.1. Актуалност на проблема и мотивация (0,5-1стр.).....	4
1.2. Цел и задачи на дипломната работа (1-2стр.).....	4
1.3. Очаквани ползи от реализацията (1-2стр.).....	5
1.4. Структура на дипломната работа (0,5-1стр.).....	5
Глава 2. Преглед на предметната област (да се замени с конкретно заглавие според заданието) (10-15стр.).....	6
2.1. Основни дефиниции.....	6
2.2. Подходи, методи (евентуално модели и стандарти) за решаване на проблемите.....	6
2.3. Съществуващи решения (практически реализации).....	6
2.4. Избор на критерии за сравнение и сравнителен анализ на решения/методи/стандарти/.....	6
2.5. Изводи.....	6
Глава 3. Използвани технологии, платформи и/или методологии (за практическото решаване на проблема) (10-15стр.).....	6
3.1. Изисквания към средствата (технологии, платформи и методологии).....	6
3.2. Видове средства (технологии, платформи и методологии) и начин и място за използването им – сравнителен анализ.....	6
3.3. Избор на средствата (технологии, платформи и методологии).....	8
3.4. Изводи.....	8
Глава 4. Анализ (10-15стр.).....	8
4.1. Концептуален модел.....	8
4.2. Потребителски (функционални) изисквания (права, роли, статуси, диаграми, ...).....	8
4.3. Качествени (нефункционални) изисквания (като напр. преносимост, използваемост, скалируемост, поддръжка, ...).....	8
4.4. Работни (бизнес) процеси.....	8
4.5. Изводи.....	8
Глава 5. Проектиране (10-15стр.).....	8
5.1. Обща архитектура – напр. слоеве, модули, блокове, компоненти.....	8
5.2. Модел на данните (напр. база данни, файлова структура, ...).....	8
5.3. Диаграми (на структура и поведение - по слоеве и модули, с извадки от кода).....	8
5.4. Потребителски интерфейс (опционално).....	8
5.5. Ресурсни и спомагателни модули (опционално).....	8
Глава 6. Реализация, тестване/експерименти и (евентуално) внедряване (10-15стр.).....	9
6.1. Реализация на модулите.....	9
6.2. Системна интеграция (опционално).....	9

6.3. Планиране на тестването - тестови сценарии, процедури,	9
6.4. Модулно и системно тестване.....	9
6.5. Анализ на резултатите от тестването и начин на отразяването им.....	9
6.6. Експериментално внедряване (технологични изисквания, инсталиране, условия, използване, ...)	9
Глава 7. Заключение (1-2стр.).....	9
7.1. Обобщение на изпълнението на началните цели.....	9
7.2. Насоки за бъдещо развитие и усъвършенстване.....	9
Използвана литература.....	9
Приложения (опционално).....	10
Приложение 1: Терминологичен речник (опционално) (1-2стр.).....	10
Приложение 2, 3,:	10
Изисквания за оформяне на дипломната работа:.....	10

Глава 1. Увод (3-5стр.)

Дипломната работа има за цел разработването на WEB базирана система, предоставяща неограничени възможности на потребителите, желаещи да трансформират разнообразен набор от текстови файлове или презентации в аудио файлове или виртуални презентации с вграден звук. Основните функции на системата са:

- Четене на HTML, TXT, PDF, DOC файлове, генериране и възпроизвеждане на аудио файлове (MP3, WAV) от прочетеното.
- Разчитане на PPT файлове и визуализирането им в Уеб вариант с вграден прочит на страниците.
- Добавяне на коментари, рейтинги и бележки към конкретен прочит.
- Предоставяне на възможност за търсене във файловете, които се намират в системата.

Продукта ще бъде разширим, позволявайки обогатяване чрез добавянето на допълнителни файлови формати, гласове за четене, начини на визуализация, възможност за командване на системата чрез глас (SpeechToText), което на практика прави възможностите му неограничени. За реализиране на системата трябва да се използват технологии и езици за програмиране, които не са платформено зависими, като така ще се премахне ограничението да се използва само конкретна операционна система, което дава допълнителна свобода на потребителите. За съхранение на системни настройки, файловете за прочит, генерираните аудио файлове и променените презентации ще се ползва база от данни. За визуализация на презентациите трябва да се обмисли използването на готово решение, което да позволява лесна интеграция и посредством което ще се улесни разработването на системата. За четене на различни видове текст ще се използват готови синтезатори на глас, който трябва да са безплатни и да работят на всички операционни системи.

1.1. Актуалност на проблема и мотивация (0,5-1стр.)

За момента съществуват няколко подобни продукта, предоставящи четене на pdf, doc, ppt. Но за всеки отделен формат трябва да инсталираме нов продукт, чуваме нови гласове, плащаме лицензи, което представлява затруднение за крайния потребител. Системата ще представлява цялостно безплатно решение, което обединява огромен набор от възможности, лесни за използване и поддръжка. Липсата на популярен и лесен за използване продукт в тази сфера, който да предоставя възможност на потребителя да го доразвие и използва според потребностите си, ме мотивира да започна работа в тази насока.

1.2. Цел и задачи на дипломната работа (1-2стр.)

Ultimate Speaker има за цел да предостави на потребителите лесен и удобен интерфейс, с който те да могат лесно да генерират прочит и визуализация на всевъзможни видове файлове. Системата може да се използва от потребители с

увреждания или просто от хора, който не могат да отделят много време за стоене пред компютъра. Запис от прочита може да бъде прослушван докато пътувате, работите, разхождате се в парка и всичко което ще ви е необходимо е обикновен MP3 плейър. Така ще се улесни възприемането му и ученето на лекции и материали ще се превърне в приятно занимание.

За да се създаде описаната система е нужно да се изпълнят следните задачи:

1. Избор на най-подходящ безплатен инструмент, позволяващ четене на текст (пр. Espeak)
2. Подбор на подходящи инструменти за организиране и разработване на проекта – език за програмиране, среда за разработка, система за компилиране и пакетиране на готовия продукт, система за съхранение на кода на приложението(source control), ОС за разработване. (пр. Eclipse, Ant, git, Java)
3. Избор на допълнителни продукти, нужни за съхранение на данни, визуализация на системата и файловете (пр. MySQL, Apache HTTP Server).
4. Изграждане на бизнес логиката на продукта и използване на софтуерни шаблони за програмиране с цел улесняване поддръжката и разширяването на системата.
5. Избор и внедряване на инструменти за разчитане на текст от HTML, PDF, DOC и PPT.
6. Тестване на системата по време на разработка (Unit Testing)
7. Подбор на готова система или изграждане на такава за визуализация на презентациите.
8. Автоматизиране инсталирането и конфигурирането на системата и използваните допълнителни продукти (MySQL, Apache, Espeak)
9. Тестване на разработеното приложение под различни операционни системи

1.3. Очаквани ползи от реализацията (1-2стр.)

1.4. Структура на дипломната работа (0,5-1стр.)

Глава 2. Преглед на предметната област (да се замени с конкретно заглавие според заданието) (10-15стр.)

2.1. Основни дефиниции

2.2. Подходи, методи (евентуално модели и стандарти) за решаване на проблемите

2.3. Съществуващи решения (практически реализации)

2.4. Избор на критерии за сравнение и сравнителен анализ на решения/методи/стандарти/...

2.5. Изводи

Глава 3. Използвани технологии, платформи и/или методологии (за практическото решаване на проблема) (10-15стр.)

3.1. Изисквания към средствата (технологии, платформи и методологии)

3.2. Видове средства (технологии, платформи и методологии) и начин и място за използването им – сравнителен анализ

За да се изберат подходящи технологии и инструменти за реализация на програмния продукт се вземат в предвид няколко основни фактора:

- Платформено независими
- Безплатни и с отворен кодовата
- Възможност за разширение

Оценявайки по тези характеристики всеки един инструмент се прави класификация и се избира най-подходящия, който в последствие ще използваме.

3.2.1 Избор на синтезатор за глас.

Развитието на двете основни технологии свързани с гласова интерпретация - *Text-to-speech* и *Speech-to-text* доведе до излизането на пазара на множество продукти предлагащи различна имплементация. Инструментите които привлякоха вниманието ми са следните:

- Синтезатор SpeechLab („Гергана“)[1]

Продуктът SpeechLab, с неговия глас „Гергана“, е първия висококачествен синтезатор на българска реч. Въпреки че към момента има още две алтернативи, гласа „Гергана“ си остава най-предпочитания от масата незрящи, особено за работа (поради добрия баланс между пъргавост и отчетливост). Като негов недостатък може да се изтъкне липсата на поддръжка на по-рядко използвани знаци от Unicode кодовата таблица, като например знакът за неравенство (\neq), и др. подобни. Освен с четенето на текстове на български, SpeechLab се справя прилично добре и с четенето на текстове на английски, което си е предимство.

SpeechLab се разпространява безплатно, но само за хората със зрителни увреждания в България. Хората със зрителни увреждания получават безплатен индивидуален лиценз за некомерсиално ползване на SpeechLab 2.0 от Фондация "Хоризонти" или от Съюза на слепите в България.

Инсталирането на **SpeechLab** е много лесно. Стартира се изпълнимия файл на инсталиращата програма, щрака се няколко пъти на бутона **Next**, след това на бутона **Install**, а накрая – на бутона **Finish** и инсталирането е завършено. Препоръчително е след края на инсталацията да рестартирате компютъра. След това гласа „Гергана“ („Gergana“) ще бъде наличен за избор в екрана за настройка на текст-към-реч в контролния панел на Windows, както и в настройките за SAPI5 в съответния екранен четец.

- **Синтезатор Innoetics Irina („Ирина“)[2]**

Гласът „Ирина“ е разработен от гръцката фирма „Innoetics“. Негово основно предимство е естествеността на звученето на гласа. Платен е (към текущия момент – 39 евро за един самостоятелен лиценз); макар и достатъчно отчетлив и приятен за слушане на текст на български, гласа „Ирина“ не се справя толкова добре с отчетливостта при четене на текстове на английски.

- **Синтезатор eSpeak (и с български глас)[3]**

Многоезичният речев синтезатор **eSpeak** е едно напълно безплатно и с отворен код решение за прочитане на текстове на глас на множество езици. Негов главен недостатък е в качеството на изходната реч – повечето потребители (особено тези свикнали с по-висококачествените речеви синтезатори), намират качеството на **eSpeak** за незадоволително. Едно от предимствата му е в пъргавостта (времето за реакция) при работа. Това обаче е за сметка на отчетливостта (разбираемостта) му. Друго негово предимство е това, че може коректно (макар и в повечето случаи – на английски) да произнася доста по-голям набор от знаци от **Unicode** кодовата таблица, в сравнение със **SpeechLab**.

Сдобиването с **eSpeak** става чрез изтегляне на инсталационния му пакет от официалния му уеб сайт: <http://espeak.sourceforge.net>, от раздела „Download“.

Наличието на Български език в **eSpeak[4]** и възможността за използването му без заплащане на лиценз наклони везните в негова посока. Архитектурата на системата позволява замяна на избрания инструмент, като единственото което трябва да се направи е да се добави имплементация за нов синтезатор на глас и

да се смени инициализирането на синтезатора при старт на системата.

3.3. Избор на средствата (технологии, платформи и методологии)

3.4. Изводи

Глава 4. Анализ (10-15стр.)

4.1. Концептуален модел

4.2. Потребителски (функционални) изисквания (права, роли, статуси, диаграми, ...)

4.3. Качествени (нефункционални) изисквания (като напр. преносимост, използваемост, скалируемост, поддръжка, ...)

4.4. Работни (бизнес) процеси

4.5. Изводи

Глава 5. Проектиране (10-15стр.)

5.1. Обща архитектура – напр. слоеве, модули, блокове, компоненти...

5.2. Модел на данните (напр. база данни, файлова структура, ...)

5.3. Диаграми (на структура и поведение - по слоеве и модули, с извадки от кода)

5.4. Потребителски интерфейс (опционално)

5.5. Ресурсни и спомагателни модули (опционално)

Глава 6. Реализация, тестване/експерименти и (евентуално) внедряване (10-15стр.)

6.1. Реализация на модулите

6.2. Системна интеграция (опционално)

6.3. Планиране на тестването - тестови сценарии, процедури, ...

6.4. Модулно и системно тестване

6.5. Анализ на резултатите от тестването и начин на отразяването им

6.6. Експериментално внедряване (технологични изисквания, инсталиране, условия, използване, ...)

Глава 7. Заключение (1-2стр.)

7.1. Обобщение на изпълнението на началните цели

7.2. Насоки за бъдещо развитие и усъвършенстване

Използвана литература

(min 10 литературни източника – статии, книги, с, форматираны съгласно MLA Style - <http://www.library.mun.ca/guides/howto/mla.php>)

[1]"Тестове: Синтезатор SpeechLab („Гергана")." Тестове: Синтезатор SpeechLab („Гергана"). N.p., 23 Apr. 2014. Web. 25 Feb. 2015. <<http://testove.nllb.bg/mod/page/view.php?id=25>>.

[2]"Тестове: Синтезатор Innoetics Irina („Ирина")." Тестове: Синтезатор Innoetics Irina („Ирина"). N.p., 23 Apr. 2014. Web. 25 Feb. 2015. <<http://testove.nllb.bg/mod/page/view.php?id=26>>.

[3]"Тестове: Синтезатор eSpeak (и с български глас)." Тестове: Синтезатор eSpeak (и с български глас). N.p., 23 Apr. 2014. Web. 25 Feb. 2015. <<http://testove.nllb.bg/mod/page/view.php?id=27>>.

[4]"ESpeak Text to Speech." ESpeak: Speech Synthesizer. N.p., n.d. Web. 01 Mar. 2015. <<http://espeak.sourceforge.net/>>

Приложения (опционално)

Приложение 1: Терминологичен речник (опционално) (1-2стр.)

Приложение 2, 3,:

напр. код или извадки от кода, наръчник на потребителя, екрани от потребителския интерфейс, спецификации, диаграми, и др.

Изисквания за оформяне на дипломната работа:

3. Чуждестранните термини да бъдат преведени, а където това не е възможно – цитирани в курсив и нечленувани.

6. Да се избягват пренасянията на нова страница на заглавия на секции, фигури и таблици.

7. Да се избягват празни участъци на страници вследствие пренасянето на фигури на нова страница.

8. Всички фигури и таблици да бъдат номерирани и именовани (непосредствено след фигурата или таблицата).

9. Всички фигури и таблици да бъдат цитирани в текста.

10. Използваните фигури от други източници да бъдат цитирани.

11. Всички цитати да бъдат отразени в списъка на използваната литература.

12. Всички източници от списъка на използваната литература да бъдат цитирани в текста.

13. Използваната литература да се цитира съгласно MLA Style - <http://www.library.mun.ca/guides/howto/mla.php>