

מסמך תכנון מפורט (Detailed Design)

מערכת לזרחי דיבור לא טיפוסי מותאם אישית בשפה העברית

1. מבוא

מסמך זה מציג את התיכון המפורט של מערכת ClearVoice. מטרתנו היא לתאר את מבנה המערכת, רכיביה העיקריים, אופן התקשרות ביניהם, תכנון בסיסי הנתונים, שימוש המשמש והאלגוריתמים המרכזיים המושגים במערכת.

המערכת נועדה לספק פתרון יישומי מותאם אישית לזרחי והמרת דיבור לא טיפוסי לפט טקסטואלי מובן, עבור אנשים על הספקturnos האוטיסטי המתמודדים עם קשיי דיבור ודיבור לא ברור.

2. סקירה כללית של המערכת

מערכת ClearVoice היא מערכת Web המיועדת לסייע לאנשים על הרצף האוטיסטי בתקשורת מיולית. המערכת מאפשרת למשתמשים להקליט דיבור דרך האתר ולקבל פלט טקסטואלי של ההקלטה.

הדגש במערכת הוא על:

- פשוטות פעולה
- שימוש ברור ולא מציף
- הציגת מיידית בצורה טקסטואלית

3. ארכיטקטורת המערכת

3.1 סגנון ארכיטקטוני

המערכת מתוכננת בארכיטקטורת Client-Server, המתאימה למערכת Web ובכוללת הפרדה ברורה בין משק המשתמש לבין הלוגיקה. גישה זו מאפשרת תחזקה קלה, הרחבת עתידית, ושיפור חווית המשתמש תוך שמירה על פשוטות ויציבות.

3.2 רכיבי המערכת העיקריים

(1) צד לקוח

אתר אינטרנט נגיש ופשוט, אחראי על:

- רישום והתחברות משתמשים
- הקלטה קול דרך הדפדפן
- שליחת ההקלטה לשרת
- הציגת הפלט הטקסטואלי של הדיבור
- ניהול מחסן הקלטות אישיות

המשק מותאם למשתמשים על הרצף האוטיסטי, עם מינימום גירויים ומבנה עקי.

(2) צד שרת

אחראי על הלוגיקה העסקית ומבצע:

- ניהול משתמשים והרשאות
- קבלת קבצי הקול מהלקוח
- הפעלת אלגוריתם תמלול דיבור לטקסט
- שמירת הקלטות ותמלולים
- החזרת פלט טקסטואלי ללקוח

(3) מודול תמלול דיבור לטקסט

רכיב ייעודי להמרת קול לטקסט, מבצע:

- קדם עיבוד בסיסי של הקול
- זיהוי דיבור
- הפקת טקסט בתוכו מתוך ההקלטה

(4) שכבת אחסון

אחראית על שמירת נתוני המשתמשים וההקלטות, כולל:

- בסיס נתונים למשתמשים
- בסיס נתונים למיחס הקולי והתמלולים

3.3 תקשורת בין רכבי המערכת

- צד הלקוח מתפרק עם צד השירות באמצעות קרייאות HTTP.
- צד השירות מעבד את ההקלטות ומפעיל את מודול התמלול.
- הפלט הטקסטואלי נשמר ומוחזק ללקוח להציג.
- כל התקשרות מתבצעת בצורה מאובטחת וברורה.

3.4 דיאגרמת ארכיטקטורה (תיאור ויזואלי)

המערכת בנויה מארבעה רכיבים עיקריים:

1. דף-דף משתמש (Frontend)
2. שרת אפליקציה (Backend)
3. מודול תמלול דיבור לטקסט
4. בסיס נתונים ו אחסון קבצים

זרימת העבודה:

- המשתמש מקליט דיבור בדף.
- ההקלטה נשלחת לשרת באמצעות קרייאות HTTP.
- השירות מפעיל את מודול התמלול ומקבל פלט טקסטואלי.
- קובץ הקול נשמר באחסון, והטקסט נשמר במסד הנתונים.
- הפלט הטקסטואלי מוחזק למשתמש להציג.

4. תכנון בסיסי הנתונים

המערכת עשויה שימוש במסד נתונים רלציוני (SQL) לצורך ניהול נתונים המשתמשים, הקלטות ותוצריו התמלול. בחירה במסד נתונים רלציוני מאפשרת שמיירה על מבנה נתונים ברור, קשרים לוגיים בין ישויות שונות, ושליפה יעילה של מידע בהתאם לצורכי המערכת. קבצי הקול עצם נשמרים כאובייקטים במסד נתונים אחסון ייעודית, כאשר במסד הנתונים נשמרים נתונים מטא דата ונתיב הגישה לקובץ.

4.1 בסיס נתונים למשתמשים

בסיס נתונים זה משמש לניהול משתמשי המערכת.

שדות עיקריים:

- מזהה משתמש (UserID)
- שם משתמש
- כתובת דואר אלקטרוני
- סיסמה
- תאריך הרשמה
- תאריך התחברות האחרונה

4.2 בסיס נתונים למחסן הקולי האישי

בסיס נתונים זה משמש לניהול הקלטות והטמלולים של המשתמשים.

שדות עיקריים:

- מזהה הקלטה (RecordingID)
- מזהה משתמש (UserID)
- נתיב לקובץ הקול
- תאריך הקלטה
- משך הקלטה
- הטקסט המתומלל
- סטטוס תמלול

קובצי הקול נשמרים כאובייקטים חיצוניים למסד הנתונים, התמלול והמידע המתאים את הקלטה נשמרים במידע טקסטואלי מוקשור.

5. תכנון ממשק המשתמש

ממשק המשתמש מתוכנן מתוך עקרונות של נגישות, פשוטות והפחחת עומס קוגניטיבי.

מאפיינים מרכזיים:

- מסכימים ברורים עם מספר פעולות מצומצם
- צבעים וגוונים טוקסט קריא
- הימנעות מהציג גרפים או מודדים מורכבים
- הצגת פلت טקסטואלי בלבד

הfonוקצינליות כוללת:

- התחברות והרשמה
- הקלטת קול
- צפיה בהקלטות קודמות
- הצגת תמלול טקסטואלי

6. אלגוריתמים ולוגיקת עיבוד עיקרית

המערכת מבצעת תהליך עיבוד מדווג של הקלטת הדיבור, הכול含 לבים עיקריים של קדם עיבוד, תמלול וניהול תוצאות.

6.1 אלגוריתם קדם עיבוד קול

שלב זה מבצע עיבוד בסיסי להקלטת המשתמש, במטרה להבין את הקלט לתהליך זיהוי הדיבור. השלב כולל:

- ניקוי רעש רך פשוט
- נרמול עצמה

מטרת שלב זה היא לשפר את איכות הקלט ולהעלות את רמת הדיווק של תהליכי התמלול.

6.2 אלגוריתם תמלול דיבור לטקסט

אלגוריתם זה אחראי על זיהוי הדיבור מתוך ההקלטה ומרתו לפلت טקסטואלי בשפה העברית.

האלגוריתם מקבל קובץ קול בקלט, מבצע עיבוד בסיסי לאות הדיבור ומפיק טקסט כתוב המיציג את תוכן ההקלטה.

השלב כולל:

- זיהוי הדיבור מתוך ההקלטה
- המרת הקול לטקסט כתוב
- טיפול בשגיאות בסיסיות (בגון קטיעים לא בחרורים או רעש רך)

ফلت של האלגוריתם הוא טקסט בלבד

6.3 ניהול תוצאות התמלול

בשלב זה מתבצע ניהול תוצרי התמלול:

- שמירת הטקסט המתמלול בסיסד הנתונים
- קישור התמלול להקלטה ולמשתמש המתאים

החזרת הפلت הטקסטואלי למשק המשתמש לצורך הצגה

7. טכנולוגיות וכלים

- **צד לקוח:**
צד הלקוח ימומש באמצעות טכנולוגיות Web סטנדרטיות (HTML, CSS, JavaScript) המאפשרות הקלטה קול דרך הדפדפן, שליחת הקלטה לשרת והציג פלט טקסטואלי למשתמש בצורה נגישה וברורה.
- **צד שרת:**
צד השרת מתוכנן להתmesh באמצעות שפת תכנות המתאימה לעיבוד נתונים, עיבוד קול ולמידת מבנה, כגון Python. השרת אחראי על ניהול משתמשים, קבלת קבצי הקול, הפעלת תהליך התמלול ושמירת הנתונים.
- **תמלול קול:**
מנגןון דיזיון דיבור והמרת דיבור לטקסט, הפועל לאחר שלב קדם עיבוד של הקלטה, ומפיק פלט טקסטואלי בלבד.
- **בסיס נתונים:**
בסיס נתונים רלציוני (SQL) לניהול משתמשים, הקלטות ותמלולים.
- **אחסון:**
קבצי Audio נשמרים כאובייקטים במערכת אחסון ייעודית, כאשר בסיס הנתונים נשמר מידע תיאורי ונתיב הגישה לקבצים.

8. סיכום

מסמן זה הציג את התקנון המפורט של מערכת ClearVoice, מערכת Web מותאמת לאנשים על הרץ האוטיסטי, המתמקדת בהמרת דיבור לא ברור לטקסט בצורה פשוטה ונגישה. התקנון מגדיש פשוטות, נגישות והפרדה ברורה בין רכיבי המערכת, ומהוות בסיס יציב למימוש הפרויקט.