

מסמך תכנון מפורט (Detailed Design)

מערכת לזיהוי 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 התקשורת ביניהם, תכנון בסיסי הנתונים, ממשק המשתמש והאלגוריתמים המרכזיים המיושמים במערכת.

המערכת נועדה לספק פתרון יישומי ומותאם אישית לזיהוי והמרת דיבור לא טיפוסי לפלט טקסטואלי מובן, עבור אנשים על הספקטרום האוטיסטי המתמודדים עם קשיי דיבור ודיבור לא ברור.

2. סקירה כללית של המערכת

ClearVoice היא מערכת Web המיועדת לסייע לאנשים על הרצף האוטיסטי בתקשורת מילולית.

המערכת מאפשרת למשתמשים להקליט דיבור דרך האתר ולקבל פלט טקסטואלי של ההקלטה.

הדגש במערכת הוא על:

- פשטות תפעול
- ממשק ברור ולא מציף
- הצגת מידע בצורה טקסטואלית

3. ארכיטקטורת המערכת

3.1 סגנון ארכיטקטוני

המערכת מתוכננת בארכיטקטורת Client-Server, המתאימה למערכת Web וכוללת הפרדה ברורה בין ממשק המשתמש לבין הלוגיקה. גישה זו מאפשרת תחזוקה קלה, הרחבה עתידית, ושיפור חוויית המשתמש תוך שמירה על פשטות ויציבות.

3.2 רכיבי המערכת העיקריים

1) צד לקוח

אתר אינטרנט נגיש ופשוט, אחראי על:

- רישום והתחברות משתמשים
- הקלטת קול דרך הדפדפן
- שליחת ההקלטה לשרת
- הצגת הפלט הטקסטואלי של הדיבור
- ניהול מחסן הקלטות אישי

הממשק מותאם למשתמשים על הרצף האוטיסטי, עם מינימום גירויים ומבנה עקבי.

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 כתוב מתוך ההקלטה

4) שכבת אחסון

אחראית על שמירת נתוני המשתמשים וההקלטות, כוללת:

- בסיס נתונים למשתמשים
- בסיס נתונים למחסן הקולי והתמלולים

3.3 תקשורת בין רכיבי המערכת

- צד הלקוח מתקשר עם צד השרת באמצעות קריאות HTTP.
- צד השרת מעבד את ההקלטות ומפעיל את מודול התמלול.
- הפלט הטקסטואלי נשמר ומוחזר ללקוח להצגה.
- כל התקשורת מתבצעת בצורה מאובטחת וברורה.

3.4 דיאגרמת ארכיטקטורה (תיאור ויזואלי)

המערכת בנויה מארבעה רכיבים עיקריים:

1. דפדפן משתמש (Frontend)
2. שרת אפליקציה (Backend)
3. מודול תמלול דיבור לטקסט
4. בסיס נתונים ואחסון קבצים

זרימת העבודה:

- המשתמש מקליט דיבור בדפדפן.
- ההקלטה נשלחת לשרת באמצעות קריאות HTTP.
- השרת מפעיל את מודול התמלול ומקבל פלט טקסטואלי.
- קובץ הקול נשמר באחסון, והטקסט נשמר במסד הנתונים.
- הפלט הטקסטואלי מוחזר למשתמש להצגה.

4. תכנון בסיסי הנתונים

המערכת עושה שימוש במסד נתונים רלציוני (SQL) לצורך ניהול נתוני המשתמשים, ההקלטות ותוצרי התמלול.

בחירה במסד נתונים רלציוני מאפשרת שמירה על מבנה נתונים ברור, קשרים לוגיים בין ישויות שונות, ושליפה יעילה של מידע בהתאם לצורכי המערכת.

קבצי הקול עצמם נשמרים באובייקטים במערכת אחסון ייעודית, כאשר במסד הנתונים נשמרים נתוני מטא דאטה 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 אחרונה

4.2 בסיס נתונים למחסן הקולי האישי

בסיס נתונים זה משמש לניהול ההקלטות והתמלולים של המשתמשים.

שדות עיקריים:

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

קבצי הקול נשמרים באובייקטים חיצוניים למסד הנתונים, התמלול והמידע המתאר את ההקלטה נשמרים כמידע טקסטואלי מקושר.

5. תכנון ממשק המשתמש

ממשק המשתמש מתוכנן מתוך עקרונות של נגישות, פשטות והפחתת עומס קוגניטיבי. מאפיינים מרכזיים:

- מסכים ברורים עם מספר פעולות מצומצם
- צבעים רגועים וטקסט קריא
- הימנעות מהצגת גרפים או מדדים מורכבים
- הצגת פלט טקסטואלי בלבד

הפונקציונליות כוללת:

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

6. אלגוריתמים ולוגיקת עיבוד עיקרית

המערכת מבצעת תהליך עיבוד מדורג של הקלטת הדיבור, הכולל שלבים עיקריים של קדם עיבוד, תמלול וניהול תוצאות.

6.1 אלגוריתם קדם עיבוד קול

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

- ניקוי רעשי רקע פשוט
 - נרמול עוצמה
- מטרת שלב זה היא לשפר את איכות הקלט ולהעלות את רמת הדיוק של תהליך התמלול.

6.2 אלגוריתם תמלול דיבור לטקסט

אלגוריתם זה אחראי על זיהוי הדיבור מתוך ההקלטה והמרתו לפלט טקסטואלי בשפה העברית.

האלגוריתם מקבל קובץ קול כקלט, מבצע עיבוד בסיסי לאות הדיבור ומפיק טקסט כתוב המייצג את תוכן ההקלטה. השלב כולל:

- זיהוי הדיבור מתוך ההקלטה
 - המרת הקול לטקסט כתוב
 - טיפול בשגיאות בסיסיות (כגון קטעים לא ברורים או רעשי רקע)
- הפלט של האלגוריתם הוא טקסט בלבד

6.3 ניהול תוצאות התמלול

בשלב זה מתבצע ניהול תוצרי התמלול:

- שמירת הטקסט המתומלל במסד הנתונים
- קישור התמלול להקלטה ולמשתמש המתאים
- החזרת הפלט הטקסטואלי לממשק המשתמש לצורך הצגה

7. טכנולוגיות וכלים

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

8. סיכום

מסמך זה הציג את התיכון המפורט של מערכת ClearVoice, מערכת Web מותאמת לאנשים על הרצף האוטיסטי, המתמקדת בהמרת דיבור לא ברור לטקסט בצורה פשוטה ונגישה. התכנון מדגיש פשטות, נגישות והפרדה ברורה בין רכיבי המערכת, ומהווה בסיס יציב למימוש הפרויקט.