

משטרת ישראל

מפרט לרכש פריטי מדף/יצור

**תוכנה לזיהוי קול ביומטרי
מפרט מס' 17/2015**

09/07/15

1. כללי :

1.1 תיאור המוצר הנדרש :

- 1.1.1 משטרת ישראל מעוניינת לרכוש תוכנה/מערכת פורנזית לזיהוי קול ביומטרי.
- 1.1.2 התוכנה תאפשר השוואה ממוחשבת (ביומטרית) בין דגימת קול של חשוד ידוע לבין דגימת קול לא ידועה (הקלטה ממדיום כלשהו) לצרכים פורנזיים (משפטיים).

2. רקע

- 2.1 ההתפתחות הטכנולוגית בעשור האחרון הובילה לפיתוח שיטות ביומטריות, וביניהן ד.נ.א, טביעות אצבע וזיהוי פנים. כחלק ממגמה זו, פותחו בשנים האחרונות שיטות ביומטריות לזיהוי דובר ואלו משולבות כיום אצל גורמי אכיפה ברחבי העולם.
- 2.2 כיום בדיקות בתחום זיהוי דובר נעשות בשיטה הפונטית בלבד (על פי שמיעה וספקטרוגרמות- שיטה קלאסית). בעתיד המטרה שהבדיקה תעשה בצורה משולבת : קלאסית + ביומטרית.

3. דרישות טכניות :

3.1 פרופיל משימה ודרישות פונקציונליות :

- 3.1.1 המערכת תופעל במז"פ במחשב נייד (בחוליית זיהוי דובר). המערכת תופעל על ידי מומחה בתיקי זיהוי דובר / השוואת קול.
- 3.1.2 ביצועים נדרשים מהמוצר :
 - השוואת דובר 1:1
 - השוואה אחד למעטים 1:n
 - המערכת תיתן התראה במידה וחומר ההשוואה לא מספיק איכותי לביצוע הבדיקה כולל קבוצות בקרה (מאגר ייחוס).
 - המערכת תתבסס על האלגוריתם העדכני ביותר הקיים בשוק לזיהוי קול בשיטה ביומטרית – I-Vector.
 - זמן בדיקה : בדיקה השוואתית (משפט מול משפט) תתבצע תוך מספר שניות (עד 15 שניות).
 - המערכת תשווה קבצי קול במספר שפות. **שפות חובה** : עברית, ערבית, רוסית, אנגלית, צרפתית.
 - במידה ונדרש מאגר קולות (מאגר ייחוס) לצורכי השוואה, הוא יסופק עם המערכת.
 - המערכת תספק עם כל המרכיבים להפעלה מידית בשפות הנ"ל.
 - המערכת תוכל לזהות דובר שהוקלט בשפות שונות (שפה מול שפה).
 - ביצועי המערכת לא יושפעו ממהירות הדיבור ומספר הדוברים.
 - למערכת תהיה יכולת סגמנטציה של שיחה על פי דובר ויצירת סינופסיס לכל דובר.
 - המערכת תוכל לעבוד עם קובצי שמע מהסוגים הבאים : WAV, mp4, mp3 לפחות.
 - במידה והמערכת תעבוד או תקבל עם פורמט אחר, יסופק פתרון להמרת קבצים.
 - המערכת תספק ריכוז תוצאות בשפה אנגלית (או/ו עברית)
 - התאמה מלאה לפעולה בסביבת מערכות הפעלה מסוג חלונות של חברת מיקרוסופט . קבלת תוצאות בדיקה בפורמט WORD או/ו PDF ו- excel
 - תוצאות הבדיקה יפורטו בדו"ח סופי בצורה Likelihood ratio ויתורגמו לסולם אורדינאלי.
 - המערכת תוכל לזהות ולהבדיל בין זכר לנקבה.
 - המערכת תוכל לתמוך לפחות בסוגי מאגרי נתונים הבאים :

- MS SQL Server 2008

- SQL Express
- MySQL 5.6
- Oracle
- EXCEL

- המערכת תוכל לעבוד במחשב מבודד או מחשב רשת בדומיין המסטרתי - ללא סביבת אינטרנט או חיבור לשרת החברה.
- המערכת תוכל לשמור "דגימת קול" מכל בדיקה, על מנת להשוות אותה בתיקים עתידיים.
- ממשק משתמש (GUI): בשפה האנגלית או עברית, ידדותי ובהיר (אפשרות להתאמה אישית על פי צרכי המשתמש).
- יכולת יצירת לוג או רישום אחר של תקלות ושגיאות של המערכת.
- **יתרון**: אפשרות לממשק API / SCRIPT להפעלת פונקציות על ידי המשתמש (גמישות של מספר השוואות ושינוי במאגרי הייחוס).
- יכולת הרכשה-אצווה (batch) של מאגר ייחוס חיצוני (של המשטרה) למערכת (בפורמט WAV).
- מפתח התוכנה לא יהיה לפי שם מחשב, תהיה אפשרות להתקין את התוכנה על מחשב אחר/אפשרות לשדרוג המחשב.

4. אספקה והתקנה:

- 4.1 עם קביעתו כזוכה, יעביר הזוכה לנציג המעבדה לראיה דיגטלית את דרישות המחשוב הנדרשות להפעלת התוכנה.
- 4.2 במידה ונדרש כרטיס מיוחד או דרבייר, יסופק ע"י הזוכה ועל חשבונו.
- 4.3 התוכנה תסופק ותותקן במחשב מסטרתי שיאופיין ע"י הזוכה במטא"ר ירושלים \ המעבדה לראיה דיגיטלית \ זיהוי דובר על ידי הזוכה, תוך 14 ימי עבודה מקבלת הזמנה מאושרת.

5. הדרכה:

- 5.1 הדרכה והשתלמויות: לפחות 3 ימים (או כל פרק זמן שימצא מתאים על ידי הזוכה ובתאום עם המשטרה) ל-2 נציגים מטעם המשטרה בארץ.
- 5.2 בכל מקרה הזוכה יישא בהוצאות הטיסה, לינה, אש"ל של נציגים מטעמו באם יידרשו להגיע לארץ לביצוע ההכשרה.
- 5.3 בסיום ההדרכה הזוכה יעביר לרשות המשטרה את חומר ההדרכה (מצגות, הדרכות, דוגמאות, חומר כתוב וכו').
- 5.4 לכל עדכון שיופץ ע"י הזוכה יצורף קובץ המפרט את השינויים והוראות ההפעלה המתאימות במידה ויש.

6. תיעוד וספרות טכנית:

- 6.1 נדרשת ספרות טכנית של המערכת ובנוסף מדריך למשתמש לגבי הפעלת המערכת ברמת המשתמש, באנגלית ו/או בעברית.

6.2 דוחות:

- 6.2.1 לתוכנה תהיה יכולת ייצור דוח ריכוז למשתמש של תוצאות בדיקות על פי חתכים שונים כגון: דו"ח שגיאות, דו"ח תוצאות מילולי, גרפי, הערכה סטטיסטית ודו"ח כמות בדיקות.

7. אחריות

- 7.1 הזוכה ייתן שירותי תמיכה טכנית ואספקת עדכוני תוכנה במדיה חיצונית, למשך שנתיים לפחות.
- 7.2 השרות יתבצע דרך אתר תמיכה מטעם הזוכה ו/או נציג תמיכה טכנית, בימים ושעות עבודה מקובלת של הזוכה.
- 7.3 הזוכה יידרש לספק מענה תוך 48 שעות מקבלת פנייה, במייל ו/או בטלפון, כולל העברת חומר הנדרש לתיקון במידה ונדרש.

8. הבטחת איכות (בדיקות איכות מקדימות):

- 8.1 משטרת ישראל מבקשת לקבל מחשב עליו מותקנת התוכנה שהוצעה במכרז, עבור ביצוע ניסוי כמפורט בנספח א'. המחשב כולל התוכנה יוגש ע"י המציעים תוך שבועיים (14 ימים) מדרישת משטרת ישראל, בסיום הבחינה יוחזרו המחשבים עם התוכנות למציעים.
- 8.2 לציין: המחשב שיסופק לבדיקה לא יחובר לרשת המשטרתית או לרשת האינטרנט.
- 8.3 במידה ונדרש מאגר קולות (מאגר ייחוס) לצורכי השוואה, הוא יסופק עם המערכת.
- 8.4 הספק רשאי להביא נציג מטעמו להפעלת המחשב והבדיקות.

שם מפרטן : רפ"ק אלן צייקובסקי

חתימת הרמ"ד האחראי על ההתקשרות : רפ"ק נטע לב-טוב שטח

תיאור בדיקת התוכנה :

1. תתבצע בדיקת התאמת הנתונים הטכניים של המערכות המוצעות לדרישות המפרט הטכני.
2. המערכת תיבדק על ידי בודקים מהמעבדה לראיה דיגיטלית במז"פ.
3. בפרמטרים הבאים כל בדיקה תרוץ פעם אחת ב-1:1 ובפעם השנייה N:1 וינוקדו :
 - 3.1. מהירות קבלת תוצאות בבדיקות שונות.
 - 3.2. דיוק המערכת בבדיקות שונות (איכויות ושפות שונות).
4. משקל כל אחת מ-16 הבדיקות שווה.
5. בכל אחת מ-16 הבדיקות, לכל מציע ינתן ניקוד יחסי בסקאלה של 1-100, כלומר המציע שייתן בבדיקת המהירות- תוצאה מהירה ביותר ובבדיקת הדיוק תוצאה מדויקת ביותר, ינוקד ב-100 נקודות, כל יתר המציעים ינוקדו בהתאם אליו באופן יחסי בהתאם לנוסחה כדלהלן :

$$\text{תוצאת מהירות / דיוק הטובה ביותר} * 100 = \text{ניקוד הבדיקה למציע} \\ \text{תוצאת מהירות / דיוק של מציע}$$

6. לאחר מכן, יסוכם לכל מציע סה"כ הנקודות שקיבל בכל אחת מ-12 הבדיקות.
7. מציע חייב לעמוד במינימום 1120 נקודות בבדיקות האיכות.
8. מציע שיקבל את הניקוד הגבוה ביותר, יקבל ציון איכות של 20. יתר המציעים ינוקדו בהתאם אליו באופן יחסי בהתאם לנוסחה כדלהלן :

$$\text{סך ניקוד המציע} * 20 = \text{ציון האיכות למציע} \\ \text{סך הניקוד הגבוה ביותר שהתקבל}$$

9. טבלה לדוגמה במקרה של 3 מציעים פוטנציאליים :

סוג בדיקה	מהירות קבלת תוצאה 1:1			דיוק התוצאה 1:1			מהירות קבלת תוצאה N:1			דיוק התוצאה N:1		
	מציע א	מציע ב	מציע ג	מציע א	מציע ב	מציע ג	מציע א	מציע ב	מציע ג	מציע א	מציע ב	מציע ג
קול זכר באיכות טובה												
קול נקבה באיכות טובה												
קול זכר באיכות נמוכה												
שיחה מעורבת של 2 דוברים (זכר ונקבה) באיכות נמוכה												
סה"כ ניקוד												

פירוט הבדיקות :

1. קול זכר באיכות טובה
2. קול נקבה באיכות טובה
3. קול זכר באיכות נמוכה
4. שיחה מעורבת של 2 דוברים (זכר ונקבה) באיכות נמוכה