# החשב הכללי - ועדת הפטור

מספר פרוטוקול **3690**

## פרטים כלליים

| המשרד הפונה | התקבל בוועדת הפטור | תאריך דיון |
|---|---|---|
| **מערך הדיגיטל הלאומי** | **15.11.2023** | **26.11.2023** |

### פרטי ההתקשרות המבוקשת

| נושא ההתקשרות | מטרת הפניה | תקנת חוק | מועד החלטת ועדת מכרזים הפונה | מספר פנייה |
|---|---|---|---|---|
| **פרויקט -קורפוס שפה** | **הארכת והרחבת התקשרות** | **3(4)(ב)(3)** | **04.09.2023** | **8001020097** |

| חברי ועדה נוכחים | תפקיד | תפקיד בוועדה |
|---|---|---|
| **אלי ביתן** | **סגן בכיר לחשב הכללי** | **יו"ר ועדה** |
| **תמר יוסף** | **יועצת משפטית** | **יועץ/ת משפטי/ת** |
| **ישי בן אלי** | **סגן בכיר לחשב הכללי** | **חבר/ת וועדה** |

## החלטת הוועדה

הוועדה ממליצה על : אישור ההתקשרות לפי תקנה 3(4)(ב)(3) לתקנות חובת המכרזים

**תקופת התקשרות** התקשרות המשך נוספת
**הערות:**
**ספקים**

| שם ספק | ח.פ. | סטטוס ספק | סוג ספק |
|---|---|---|---|
| האקדמיה ללשון העברית | 500500160 | | |

**פרטי תקופה**

| תאריך תחילה מבוקש | תאריך תחילה מאושר | תאריך סיום מבוקש | תאריך סיום מאושר | אומדן | סכום מאושר | מטבע |
|---|---|---|---|---|---|---|
| 22.11.2023 | 26.11.2023 | 31.12.2025 | 31.12.2025 | 2,100,000 | 2,100,000 | שקלים ישראליים חדשים |

ההחלטה מותנית באישור של: החשב הכללי

## דיון ונימוקים

1. מערך הדיגיטל הלאומי (להלן - **המערך**) מבקש התקשרות המשך עם האקדמיה ללשון העברית (להלן - **האקדמיה**) לטובת השלמת תהליך התיוג.
2. ביום 03.11.2020, אישרה ועדת הפטור המשרדית התקשרות עם האקדמיה לתקופה שמיום 03.11.2020 ועד ליום 31.12.2023 בהיקף של 2,100,000 ₪ מכוח תקנה 3(7) לתקנות חובת המכרזים, לצורך הקמת פרויקט קורפוס שפה. הפרויקט נועד על מנת ליצור אלגוריתם אשר ידע להשתמש בשפת העברית בצורה מלאה ותקנית, והוא מיועד לשימוש עבור משרדי הממשלה השונים.
3. על פי ההסכם שנחתם, המערך משלם לאקדמיה לפי מספר מילים מתויגות.
4. נכון לשלב הנוכחי, עד כה פורסמו באתר data.gov.il חמש מנות הראשונות של הטקסטים שכוללות כ- 95,000 מילים, וכ-

35,000 מילים נוספות נמצאות בשלבים אחרונים של הכנה לפני פרסום ב data.gov.il או להעברה למשרד הבריאות ומשרד הרווחה בהתאמה, שהם אחראים במקרים אלו על טיפול בטקסטים, עקב היבטי הפרטיות.

5. פרסום קבצי התיוג הנותרים ימשיך להיעשות עד לסוף שנת 2023 תוך השלמה בשנת 2024 במידת הצורך, לאחר השלמת שליפת הטקסטים המוכנים מתוך מערכות האקדמיה והמרתם לקבצי פלט בהתאם לתקן שגובש.

6. בתהליך היו עיכובים שונים שנבעו ממורכבות הכנת התקן, מאתגרים פנימיים של האקדמיה הן מול חברת התיוג והן פנימית, ומהצורך בתיקונים בתוכנת התיוג עקב כך. כמו כן, היו עיכובים בצד לקוחות התיוג – בפרט במשרד הרווחה, עקב הצורך בביצוע עילום (התממה) של הטקסטים שנועדו לתיוג.

7. הערכת מצב שביצע המערך לגבי המשך תהליך התיוג, על מנת לוודא את הצורך של הלקוחות הנוספים ולקוחות פוטנציאליים חדשים מעלה כי קיימת התעניינות אצל מספר גורמים, אולם יש לקחת בחשבון שלא בוצע תהליך מלא של שימוש בטקסטים לצורך ייצור יישומי בינה מלאכותית באופן מלא, שיאפשר בחינה של רמת התועלת של שימושים אלו, תוך שיפור יישומי עיבוד שפה טבעית (processing language natural – NLP) .

8. לפיכך, האסטרטגיה שהיחידה המקצועית גיבשה היא לפעול בשני שלבים כך שבשלב ראשון יושלם תיוג מילים להשלמת הפעילות על פי המידע שקיים בידי היחידה המקצועית ומאפשר ביצוע מידי.

9. לאור האמור, מבקש המערך להרחיב את ההתקשרות עבור הגדלת היקף התיוג של הפרויקט בהיקף של 300,000 מילים נוספות (לפי 7 ₪ למילה), על מנת לאפשר למשרדים נוספים להצטרף למיזם הקורפוס ולאפשר לגופים (ממשלתיים וחוץ ממשלתיים) העוסקים בפיתוח אלגוריתמים ויישומי NLP שיוכלו לספק שירותים מגוונים על בסיס קורפוס שפה עברית. יודגש, כי המערך ימשיך בבחינה מול המשרדים של רמת התועלת לגבי הרחבת הפרויקט, בכדי להעריך את היקף כמות התיוג הנדרשת בעתיד.

10. בהתאם לתקנה 3(7) לתקנות חובת המכרזים ההתקשרות מבוצעת בהתאם לסעיף 2 לחוק האקדמיה "תפקיד המוסד הוא לכוון את התפתחות הלשון העברית על יסוד חקר הלשון לתקופותיה ולענפיה". על פי המערך, האקדמיה היא המוסד הפוסק בתחום הלשון העברית, והחלטותיה בענייני דקדוק, כתיב, מינוח ותעתיק אשר מחייבות את רשויות המדינה. האקדמיה מתמחה בתיוג טקסטים עבריים וברשותה קורפוס השפה העברית המתויג ידנית הגדול בעולם. בנוסף האקדמיה כגוף סטטוטורי יכולה להבטיח רצף של טיפול בקורפוס לאורך זמן. רציפות שאינה יכולה להיות מובטחת בהתקשרות עם גופים אקדמאיים או גופים מסחריים.

11. התקשרות ההמשך בתנאים זהים להתקשרות המקורית.

12. מר בנימין נדב, רפרנט משרדים מנהליים ממליץ לאשר את התקשרות ההמשך.

**מאשרים**

| פרטי אישור | שם פרטי ושם משפחה | תפקיד | תפקיד בוועדה | הערות |
|---|---|---|---|---|
| אישור | יהלי רוטנברג | החשב הכללי באוצר | גורם מאשר נוסף | |
| אישור | אלי ביתן | סגן בכיר לחשב הכללי | יו"ר ועדה | |
| אישור | ישי בן אלי | סגן בכיר לחשב הכללי | חבר/ת וועדה | |
| אישור | תמר יוסף | יועצת משפטית | יועץ/ת משפטי/ת | |

אישור החשב הכללי:

אני מחליט על אישור ההתקשרות כמפורט בהחלטה מספר 3000009659 לעיל,

לפי תקנה 3(4)(ב)(3) לתקנות חובת המכרזים.



נחתם דיגיטלית
12.12.2023