

מנוע מודיעין פטנטים גלובלי (GPIE)

מסגרת לחיפוש פטנטים לאומי ובינלאומי, סיווג, מיפוי משפחות, וניתוח חפיפת תביעות

מחבר:

כבוד טירי ג' מייסון, מנהל הלשכה לניתוח חישובי

סוג המסגרת:

אלגוריתם מודיעין פטנטים ואיתור אמנויות קודמות (Prior-Art)

היקף ותפקיד:

מערכות לאומיות, אזוריות ובינלאומיות | גילוי, נורמליזציה, סיווג, השוואה ודירוג אנליטי

1. הגדרת המערכת

מנוע מודיעין פטנטים גלובלי (GPIE) הוא מסגרת חישובית שנועדה לבצע חיפוש פטנטים שיטתיים על פני מאגרי מידע קנייניים לאומיים, אזוריים ובינלאומיים. במקום להתייחס לחיפוש פטנטים כשאלת טקסט קונבנציונלית, GPIE ממפה את החיפוש כבעיית אחזור מידע מבוססת גרף:

$$P_{\text{search}} = \Phi(K, C, F, J, T, Q, R)$$

כאשר:

- K = מילות מפתח ומונחים טכניים
- C = מבני סיווג פטנטים
- F = יחסי משפחת פטנטים
- J = מידע שיפוטי/גאוגרפי
- T = יחסים זמניים (טמפורליים)
- Q = תביעות ומגבלות טכניות
- R = מדדי רלוונטיות ודמיון

מטרה מרכזית: מטרה אינה רק למצוא פטנטים המכילים מילים דומות, אלא לקבוע: אילו המצאות שפורסמו בעבר, בקשות פטנט, משפחות פטנטים, תביעות, פרסומים וגילויים טכניים רלוונטיים מהותית להמצאת הנושא, באיזה תחומי שיפוט, באילו תקופות, ובאיזו מידה של חפיפה טכנית ומשפטית?

2. ארכיטקטורת הליבה

מורכב מתשע שכבות עיקריות GPIE:

$GPIE = L_1 + L_2 + L_3 + L_4 + L_5 + L_6 + L_7 + L_8 + L_9$

שכבה	שם	פעולה עיקרית
שכבה 1	פירוק המצאה (Invention Decomposition)	המרת ההמצאה המוצעת לאלמנטים טכניים בדידים.
שכבה 2	הרחבת מינוח (Terminology Expansion)	הפקת מילים נרדפות, מינוח היסטורי, קיצורים, מינוח מדעי ותיאורים חלופיים.
שכבה 3	גילוי סיווג (Classification Discovery)	זיהוי מערכות סיווג רלוונטיות כגון CPC, IPC, USPC, Locarno ואחרות.
שכבה 4	איתור גלובלי (Global Retrieval)	חיפוש במאגרי פטנטים לאומיים, אזוריים ובינלאומיים.
שכבה 5	שחזור משפחת פטנטים (Family Reconstruction)	קביעה האם מסמכים הנראים נפרדים מייצגים את אותה ההמצאה הבסיסית.
שכבה 6	מודיעין תביעות (Claim Intelligence)	חילוץ תביעות עצמאיות ותלויות ופירוקן למגבלות טכניות.
שכבה 7	ניתוח זמני של אמנות קודמות (Temporal Analysis)	קביעת תאריכי פרסום, תאריכי בכורה, תאריכי הגשה, המשכים, חצאיות, מענקים וסטטוס פקיעה.
שכבה 8	ניתוח דמיון וחפיפה (Similarity & Overlap)	השוואת ההמצאה המוצעת מול גילויים מזוהים.
שכבה 9	פלט אנליטי-משפטי (Legal-Analytical Output)	הפקת נוף פטנטים מודע-שיפוט וזיהוי בעיות חידוש, חפיפה וחופש פעולה (FTO).

3. מנוע פירוק המצאות

הפעולה הראשונה לעולם אינה צריכה להיות חיפוש מילות מפתח. תחילה יש לייצג את ההמצאה כוקטור מובנה:

$I = \{S, F, M, E, C, P, O\}$

כאשר: **S** = רכיבי מערכת, **F** = יחסים פונקציונליים, **M** = מנגנונים, **E** = קלטים/פלט, **C** = ארכיטקטורת בקרה, **P** = שלבי תהליך, ו-**O** = תוצאה תפעולית מיועדת.

דוגמה מערכת אנרגיה מבוססת בינה מלאכותית עשויה להתפרק ל: תת-מערכת קלט אנרגיה, תת-מערכת חישה, תת-מערכת חישובית, אלגוריתם בקרה, מנגנון המרה, תת-מערכת אגירת אנרגיה, מנגנון משוב ותת-מערכת פלט. פירוק זה מונע ממנוע החיפוש להפוך לתלוי במונחים הספציפיים של הממציא.

4. מחולל אונטולוגיה טכנית

יוצר גרף מינוחים $G_T = (V, E)$ שבו כל צומת מייצג מושג טכני וכל קשת מייצגת קשר סמנטי. עבור כל אלמנט טכני, המנוע מייצר: מינוח מדויק, מילים נרדפות, קיצורים, ראשי תיבות, מינוח היסטורי, מינוח הנדסי, מינוח מדעי, איות חלופי, מינוח ברמת רכיב, תיאורים פונקציונליים, תיאורים מוכונו-תוצאה, מושגים רחבים ומושגים צרים.

לדוגמה: קוונטי" עשוי להתרחב למושגים הכוללים: חישה קוונטית, מדידה קוונטית, גלאי קוונטי, חישה משופרת-קוונטית, מטרולוגיה קוונטית, חיישן אטומי, חיישן מוליך-על, חישה מבוססת-ספין וגלאי מבוסס-פוטון. מנוע ההרחבה צריך להבחין בין שווי ערך סמנטיים אמיתיים לבין מושגים קשורים בלבד.

5. הפקת וקטור חיפוש & 6. היררכיית חיפוש

המערכת מייצרת מספר וקטורי חיפוש עצמאיים במקום שאילתת בוליאנית ענקית אחת:

$$V = [V_K, V_C, V_F, V_Q, V_T, V_H]$$

(מילות מפתח, סיווג, פונקציונלי, שפת תביעות, טמפורלי, סמנטי היברידי)

מבצע חיפושים ב-9 שכבות הולכות ומתרחבות GPIE

1. **מינוח מדויק:** חיפוש ביטויים טכניים מדויקים.
2. **הרחבת נרדפות:** החלפת מונחים בשווי ערך מוכרים.
3. **חיפוש פונקציונלי:** חיפוש אחר מה שההמצאה עושה, ולא אחר מה שהיא נקראת.
4. **חיפוש רכיבים:** חיפוש תת-מערכות בודדות באופן עצמאי.
5. **חיפוש סיווג:** חיפוש במבני CPC/IPC ומבני סיווג אחרים.
6. **חיפוש ציטוטים:** מעקב אחר ציטוטים לאחר, קדימה, של בוחן, של מגיש, וציטוטי NPL.
7. **חיפוש משפחות:** הרחבת מסמכים שהתגלו למשפחות הפטנטים המלאות שלהם.
8. **חיפוש ממציא/מקבל העברה:** חיפוש עבודה קשורה לפי ממציא, מגיש, מקבל העברה, חברת בת, קודם או יורש.
9. **דמיון סמנטי:** שימוש בשיטות הטבעה/אחזור סמנטי לאיתור גילויים דומים מבחינה טכנית בעלי אוצר מילים שונה.

7. מטריצת שיפוט גלובלית

כל מסמך שאותר משויך לצומת שיפוט: $J_i = (country, region, office, status)$. המערכת מבחינה בין:

- **משרדים לאומיים:** ארצות הברית, קנדה, בריטניה, גרמניה, צרפת, איטליה, ספרד, יפן, סין, דרום קוריאה, הודו, אוסטרליה, ברזיל, מקסיקו, ישראל, סינגפור ומשרדים לאומיים אחרים.
- **מערכות אזוריות:** ארגוני פטנטים אזוריים רלוונטיים (למשל, EPO, EA, ARIPO, OAPI).

• **המערכת הבינלאומית:** שכבת הפרסום של PCT/WIPO מטופלת כשכבת ניתוב גלובלית ולא כמדווחת אוטומטית כפטנט מורשה עולמי.

8. שחזור משפחת פטנטים

קביעה האם מסמכים מרובים מייצגים את אותה ההמצאה הבסיסית באמצעות גרף משפחה $G_F = (P, E)$ (כאשר P = מסמכי פטנט, E יחסי בכורה/המשך/חלוקה/שלב לאומי). המערכת מזהה: מסמך בכורה מוקדם ביותר, שרשרת בכורה, בקשת PCT, בקשות שלב לאומי, המשכים, CIPs, חלוקות, מענקים, בקשות נטושות, פטנטים שפגו ובקשות תלויות ועומדות. זה מונע מהמצאה יחידה להיספר בטעות כעשרות הפניות אמנות קודמות עצמאיות.

9. פירוק תביעות & 10. מטריצת חפיפת תביעות

$$C_i = \{L_1, L_2, \dots, L_n\}$$

תביעות פטנט מומרות למגבלות מובנות. עבור תביעה עצמאית: כל מגבלה ממוינת כמבנית, פונקציונלית, יחסית, זמנית, מספרית, חומרית, אלגוריתמית, פרוצדורלית, סביבתית או קשורה לבקרה.

"מערכת הכוללת A :" המוגדרת לתקשר עם B כדי לייצר "C" הופכת ל- $C = \{A, B, A \rightarrow B, f(A,B)=C\}$. המנוע משווה מגבלות ולא משפטים שלמים.

עבור פטנט מועמד P והמצאה מוצעת I , חפיפת התביעות מחושבת כדלקמן:

$$O(I, P) = \frac{\sum_{i=1}^n (w_i M_i)}{\sum_{i=1}^n w_i}$$

כאשר $M_i = 1$, $M_i \in [0, 1]$ כאשר מגבלה i נחשפה מהותית ($M_i = 0$ אם אינה קיימת), ו- w_i הוא משקל החשיבות.

11. דירוג רלוונטיות

כל מסמך שאותר מקבל ציון רלוונטיות מורכב:

$$R(P) = w_k K + w_c C + w_s S + w_q Q + w_f F + w_t T + w_x X$$

כאשר המשקלים ניתנים להתאמה לפי מצב חיפוש:

• **מצב אמנות קודמת (Prior-Art):** הגדלת w_q, w_f

• **מצב נוף פטנטים:** הגדלת w_c, w_s, w_f

12. מודיעין זמני, 13. גרף ציטוטים, & 14. מודיעין סמנטי

מודיעין זמני (Temporal) מר על ארבעה תאריכים נפרדים $D = \{D_p, D_f, D_{pub}, D_{app}\}$ (בכורה, הגשה, פרסום, מענק). המנוע לעולם אינו מתייחס לתאריכים אלו כאל הניתנים להחלפה.

הרחבת גרף ציטוטים: חוקר באופן רקורסיבי עד שקצב הגילוי השולי יורד מתחת לסף ϵ :

$$P_0 \rightarrow C_{backward} \rightarrow C_{forward} \rightarrow F_{family} \rightarrow I_{inventor} \rightarrow A_{assignee}$$

$$(\Delta N_{relevant} / \Delta N_{searched}) < \epsilon$$

שכבת חיפוש סמנטי מתווספת על ידי אחזור וקטורי. הטבעות $e_P, e_I \in \mathbb{R}^n$ מושוות באמצעות דמיון קוסינוס: $S(P, I) = \cos(e_P, e_I)$ משמש לאחזור בלבד; מסקנות משפטיות דורשות ניתוח ברמת תביעות וספציפי לתחום שיפוט.

15. ספרות שאינה פטנט (NPL), 16. סטטוס, & 17. ביטחון חיפוש

ספרות שאינה פטנט (NPL): כוללת מאמרים מדעיים, דיוני כנסים, עבודות דוקטורט, מדריכים טכניים, תקנים, אתרי אינטרנט בארכיון, פרסומים הנדסיים, מסמכים רגולטוריים והדגמות ציבוריות. ממיון כמקורות {Patent, Application, NPL, Product, Standard, Archive, Other}.

מנוע סטטוס שיפוט: מצב מנומל P_{Status} (in) {Pending, Granted, Abandoned, Expired, Lapsed, Revoked, Withdrawn, Unknown} המקורי על סטטוס המשרד המקורי.

ציון ביטחון חיפוש: $f_{Conf} =$ (איכות מקור, כיסוי תביעות, ודאות משפחה, ודאות זמנית, ודאות שיפוטית).

טווח ציונים	פרשנות
0.90 - 1.00	ביטחון גבוה מאוד
0.75 - 0.89	ביטחון גבוה
0.50 - 0.74	ביטחון בינוני
0.25 - 0.49	ביטחון נמוך
0.00 - 0.24	ראיות בלתי מספקות

18. מכונת מצבי חיפוש

מצבים	כרצף	פועל	האלגוריתם
עם $S_0 \rightarrow S_1 \rightarrow S_2 \rightarrow S_3 \rightarrow S_4 \rightarrow S_5 \rightarrow S_6 \rightarrow S_7$			
לולאת משוב $S_5 \rightarrow S_2$ לשילוב מינוח שהתגלה לאחרונה בשלבי חיפוש מוקדמים יותר.			

19. אלגוריתם הליבה

אלגוריתם ביצוע GPIE

קלט: תיאור הממצא | תאריך רלוונטי | תחומי שיפוט יעד | מטרת חיפוש

1. ניתוח ההמצאה.
2. פירוק ההמצאה לאלמנטים טכניים.
3. יצירת אונטולוגיה טכנית.
4. יצירת מילים נרדפות ותיאורים פונקציונליים.
5. זיהוי CPC/IPC וקודי סיווג רלוונטיים.
6. ביצוע חיפוש מילות מפתח מדויקות.
7. ביצוע חיפוש מילות מפתח מורחבות.
8. ביצוע חיפוש סיווג.
9. ביצוע חיפוש ממוקדי תביעות.
10. דירוג תוצאות ראשוניות.
11. הרחבת משפחות פטנטים בעלות רלוונטיות גבוהה.
12. מעבר על ציטוטים לאחר.
13. מעבר על ציטוטים קדימה.
14. חיפוש ממציאים ומקבלי העברה.
15. חיפוש ספרות שאינה פטנט קשורה.
16. חילוץ תביעות ומגבלות טכניות.
17. נורמליזציה של תאריכים.
18. נורמליזציה של סטטוס שיפוט.
19. בניית גרף משפחת פטנטים.
20. בניית גרף ציטוטים.
21. חישוב דמיון סמנטי.
22. חישוב חפיפת תביעות טכניות.
23. החלת מסננים זמניים.
24. הסרת כפילויות של חברי משפחה.
25. דירוג תוצאות.

26.	זיהוי פערחי חיפוש.
27.	חזרה על חיפושים עבור פערים בלתי פתורים.
28.	עצירה כאשר הגילוי השולי יורד מתחת לסף.
29.	הפקת נוף פטנטים.
30.	הפקת מטריצת אמנות קודמות.
31.	הפקת מטריצת חפיפת תביעות.
32.	הפקת מטריצה שיפוטית.
33.	הפקת ציוני ביטחון.
34.	הפקת דוח אנליטי סופי.
פלט: דוח מודיעין פטנטים גלובלי	

20. מטריצות פלט עיקריות

א. מטריצת נוף פטנטים (Patent Landscape Matrix)

משפחת פטנטים	תחום שיפוט	בכורה	פרסום	סטטוס	רלוונטיות טכנית	רלוונטיות תביעות
FAM-001	US, EP, JP	2018-04-12	2019-10-18	מורשה (Granted)	גבוהה	גבוהה
FAM-002	WO, CN	2020-01-15	2021-07-22	בהמתנה (Pending)	בינונית	נמוכה

ב. מטריצת אמנות קודמות (Prior-Art Matrix)

מזהה (Reference)	גילוי	תאריך רלוונטי	חפיפה טכנית	חפיפת תביעות	ביטחון
US10123456B2	בקרת תת-מערכת A + B	2017-09-01	85%	70%	0.92
NPL-IEEE-2019	אלגוריתם משוב חיישן	2019-02-10	60%	לא ישים (N/A)	0.85

ג. מטריצת משפחה (Family Matrix)

בכורה שורש	PCT	US	EP	CN	JP	KR
US201862123456	WO2019123456	US10987654	EP3654321	CN110987654	JP6876543	KR10202012345

ד. מטריצת מגבלות תביעות (Claim-Limitation Matrix)

מגבלה	המצאה מוצעת	פטנט A	פטנט B	פטנט C
A רכיב מבני L1:	✓			✓
A ↔ B צימוד פונקציונלי L2:	✓		✓	
אלגוריתם סף בקרה L3:	✓	✓		✓

21. מצבי מטרת חיפוש

- **מצב 01 - חיפוש חידוש (NOVELTY SEARCH):** שאלה עיקרית: האם נושא התביעה כבר פורסם?
- **מצב 02 - חיפוש התקדמות המצאתית / מובנות מאליה (INVENTIVE-STEP / OBVIOUSNESS):** שאלה עיקרית: האם שילובי גילויים קיימים מתכנסים מהותית על ההמצאה המוצעת?
- **מצב 03 - סקר חופש פעולה (FREEDOM-TO-OPERATE SCREEN):** שאלה עיקרית: אילו זכויות אכיפות פוטנציאליות עשויות להצטלב מהותית עם הפעילות המתוכננת?
- **מצב 04 - נוף פטנטביליות (PATENTABILITY LANDSCAPE):** שאלה עיקרית: איזה שטח טכני נותר פתוח יחסית?
- **מצב 05 - מודיעין מתחרים (COMPETITOR INTELLIGENCE):** שאלה עיקרית: אילו ארגונים וממציאים מפתחים טכנולוגיה בתחום זה?
- **מצב 06 - נוף טכנולוגי (TECHNOLOGY LANDSCAPE):** שאלה עיקרית: כיצד תחום טכנולוגי שלם מבוזר על פני פטנטים, משפחות, תחומי שיפוט וארגונים?
- **מצב 07 - ניתוח תיקים (PORTFOLIO ANALYSIS):** שאלה עיקרית: כיצד משפחות פטנטים בודדות קשורות בתוך תיק הקניין הרוחני של ארגון?

22. מדד שלמות & 23. ארכיטקטורה אנטי-הטיה

מדד שלמות (Completeness Metric): חיפוש אינו מוגדר כ"שלם" רק משום שמספר מוגדר של מאגרי מידע נחקרו. השלמות מחושבת כ- $\text{Gamma} = 1 - (N_{\text{new}} / N_{\text{examined}})$. ככל ש- $\text{Gamma} \rightarrow 1$, החיפוש מתקרב לרוויה. כיסוי המערכת ניכר דרך $\text{Coverage} = (J, C, T, S, F, N)$.

ארכיטקטורה אנטי-הטיה (Anti-Bias Architecture): נתיבי חיפוש עצמאיים, הרחבה עיוורת, הסרת כפילויות של משפחות, הרחבת ציטוטים ותיעוד חיפושיים שליליים (היסטוריית חיפוש ניתנת לביקורת).

24. שביל ביקורת & 25. ארכיטקטורת מודיעין סופית

שביל מייצב ביקורת רשומות (Audit Trail) :
 $A_i = \{ \text{timestamp, query, database, jurisdiction, filters, results, ranking, analyst/system, decision} \}$

ארכיטקטורת מודיעין סופית:

$$GPIE = D + O + C + R + F + Q + T + J + N + A$$

[Invention Ontology] → [Classification] → [Global Retrieval] → [Patent Graph] → [Claim Graph] → [Temporal Analysis] → [Jurisdictional Analysis] → [Patent Intelligence]

26. עקרון מנחה

עקרון מנחה: הפש את ההמצאה לפי מה שהיא נקראת, מה שהיא עושה, מה שהיא מכילה, כיצד הרכיבים שלה פועלים הדדית, כיצד התביעות שלה בנויות, היכן מופיעים האבות הטכנולוגיים שלה, וכיצד גילויים אלו מתפשטים דרך משפחות פטנטים גלובליות.