



How Does UML Look and Sound?

Using AI to Interpret UML Diagrams through Multimodal Evidence

Noor Hatoom & Shihriban Khatib

20 May 2025



הקדמה

**דיאגרמות UML הן כלי מרכזי בלמידה ובפיתוח
של מערכות מידע, הן עוזרות לנו לייצג מבנה
של מערכת, תהליכיים, ואינטראקציות.**

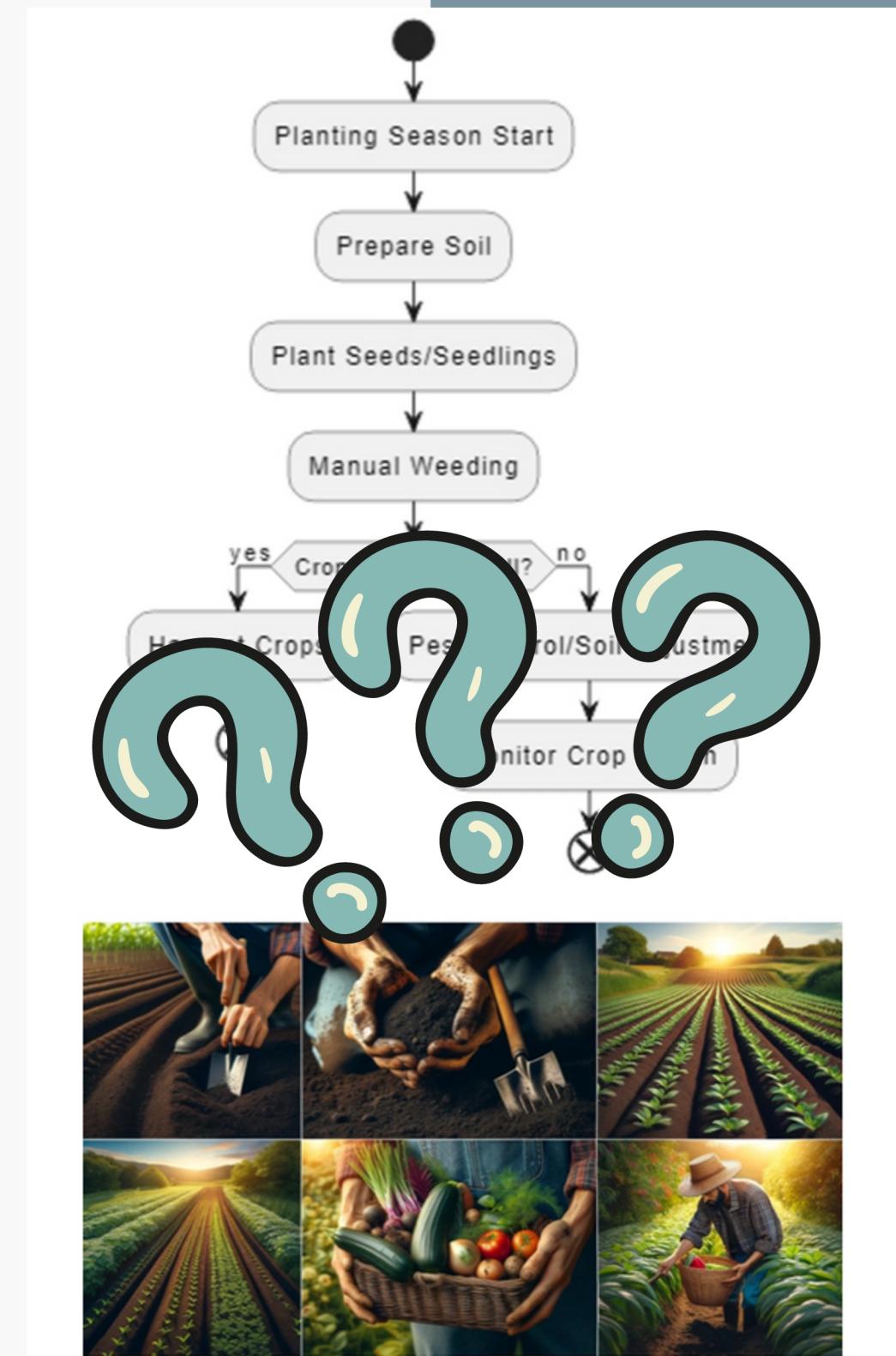
הרבה סטודנטים ובכלל אנשי מקצוע בתחום הדרך מתקשם להבין
ולהתחבר לדיאגרמות האלה הן מופשטות, סטטיות ולייטיים נראות
"מנוטקות" מהעולם האמיתי.



שאלת המחקר

RQ: How can generative AI enhance the comprehension and creation of UML diagrams by providing contextual real-world examples?

כיצד בינה מלאכותית גנרטיבית יכולה לשפר את הבנה והיצירה של דיאגרמות UML על ידי מתן דוגמאות קונטקטואליות מהעולם האמיתי?



הפתרון

ModSense

שיטת modSense מציעה למידה רב-חושית לדיאגרמות UML, באמצעות תМОנות וסבירים קוליים שנוצרים על ידי בינה מלאכותית. המערכת מתאימה את הרmachה לסגנון הלמידה של כל סטודנט בין אם הוא חזותי או שמיוני ומחברת בין הדיאגרמות לעולם האמתי. כך היא מאפשרת למידה נגישה, מותאמת אישית ועמוקה יותר של מודלים מושגים.



הפתרונות – ModSense



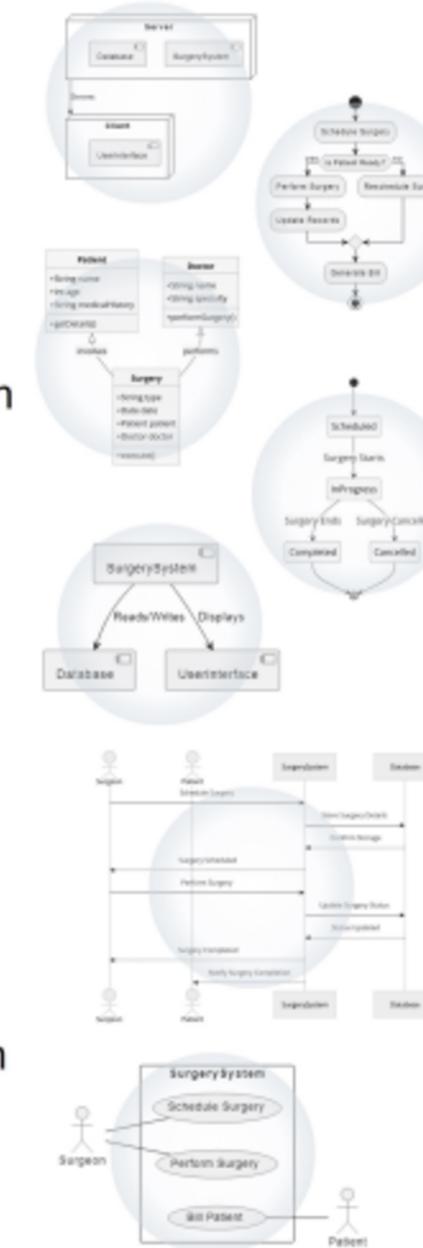
מה עושה לפרש את שני סוגי הדיאגרמות (מבנהו וдинמיות) באופן חזותי וקולני?

4 A. Gavric et al.

(UML diagrams)

- ① Class Diagram
- ② Object (Instance) Diagram
- ③ Component Diagram
- ④ Composite Structure Diagram
- ⑤ Package Diagram
- ⑥ Deployment Diagram
- ⑦ Profile Diagram

Structural UML diagrams



- ① Use-Case Diagram
- ② Sequence Diagram
- ③ Activity Diagram
- ④ State Diagram
- ⑤ Communication Diagram
- ⑥ Timing Diagram
- ⑦ Interaction Overview Diagram

Behavioral UML diagrams

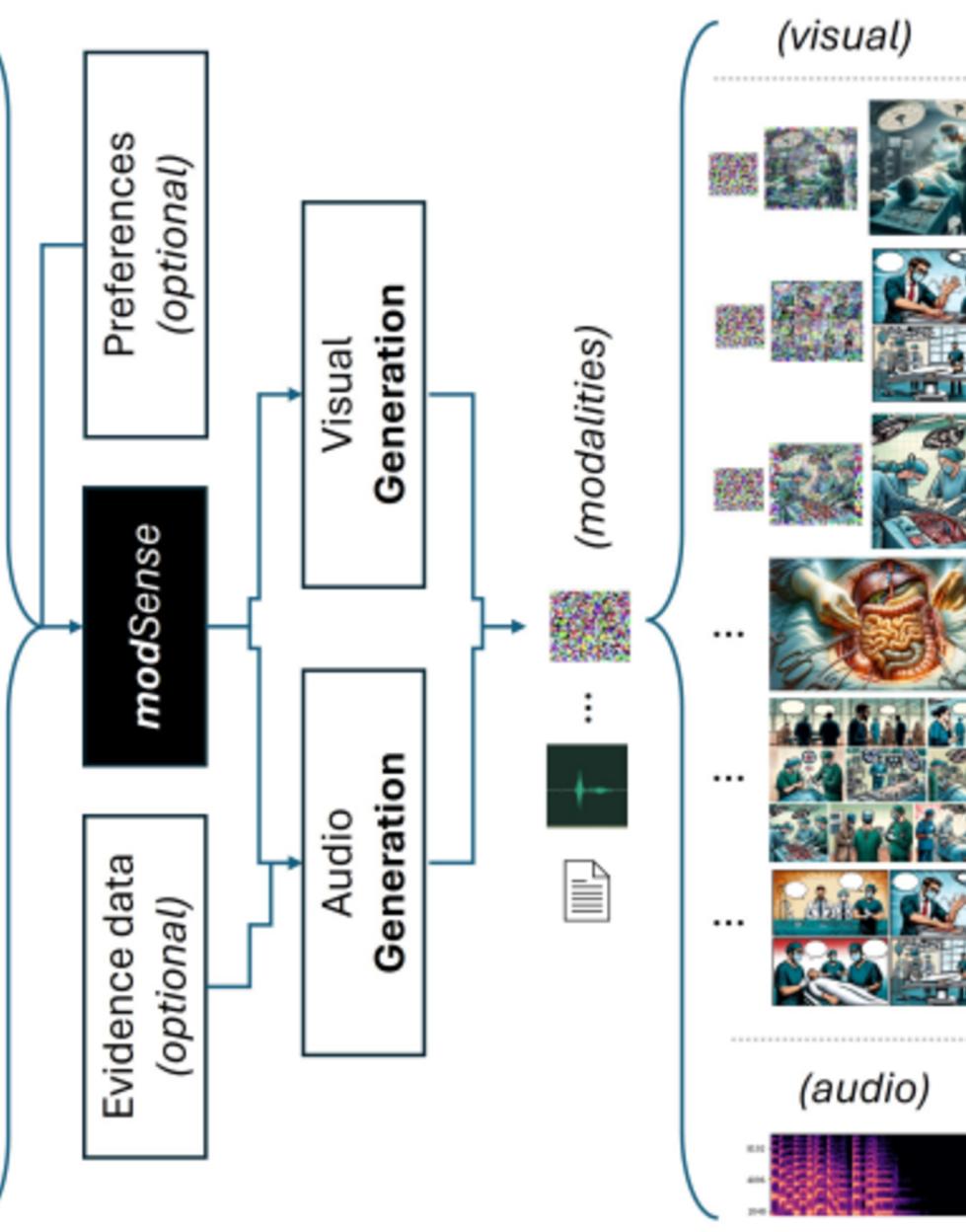
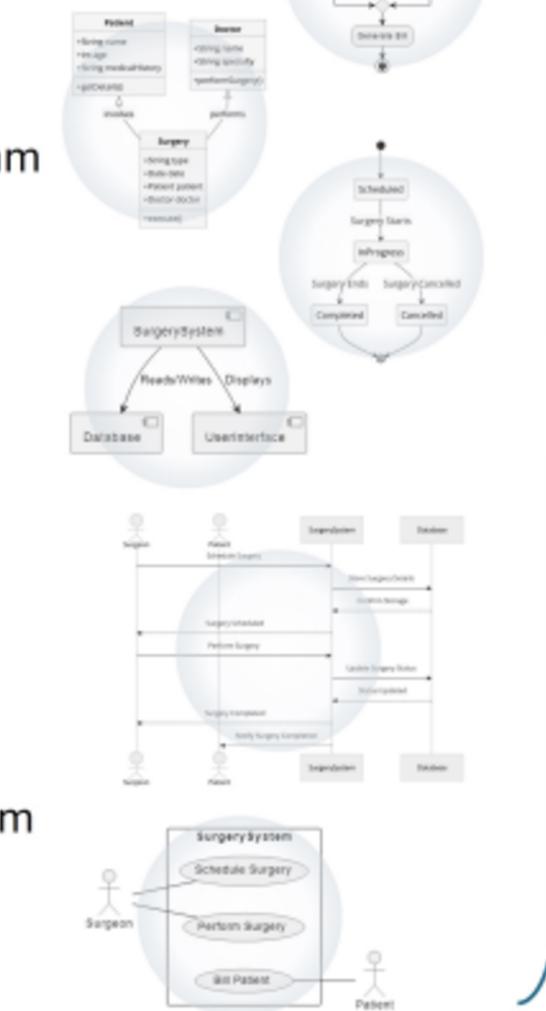
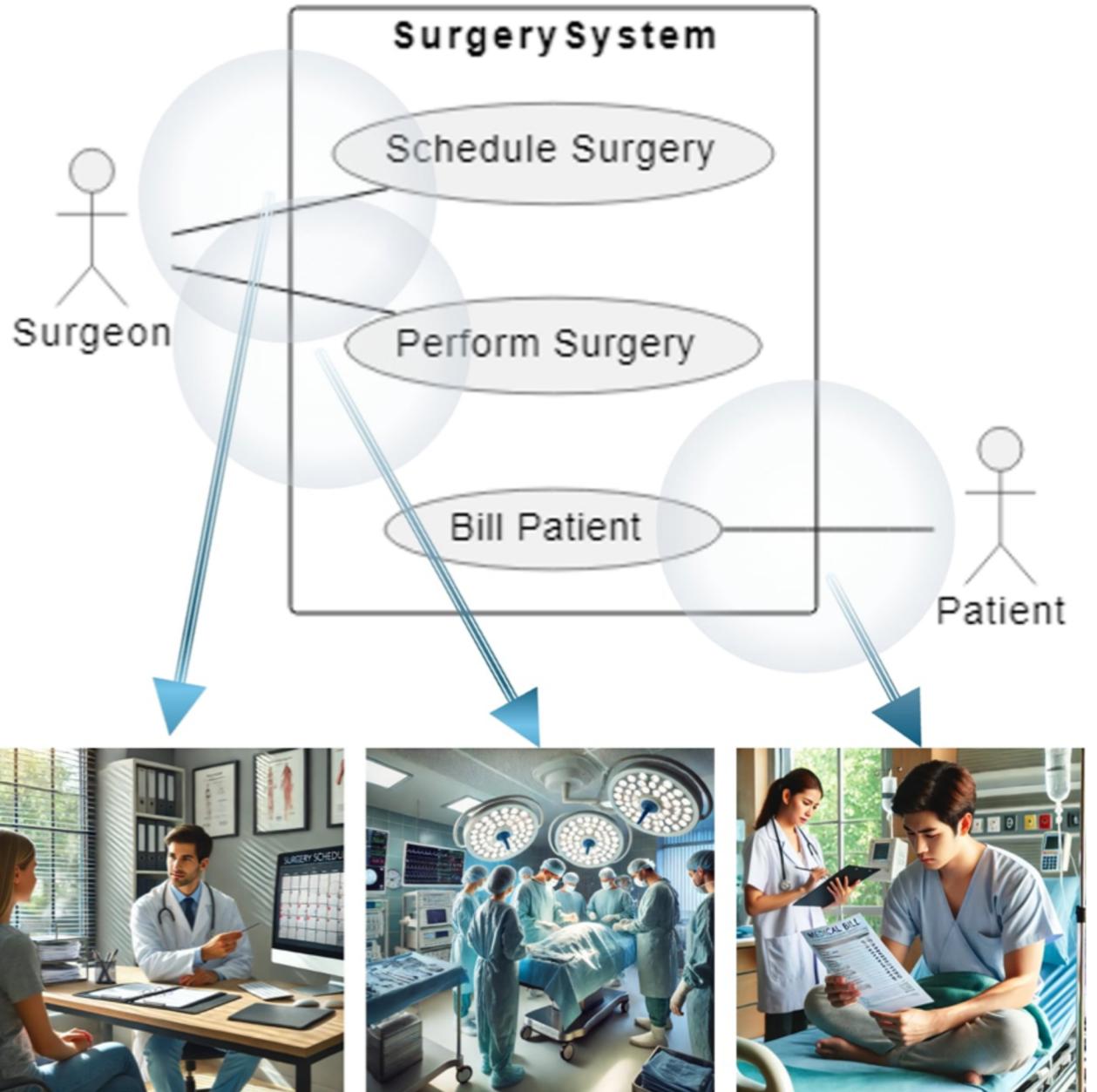


Fig. 1. *modSense* platform overview.



יכולות ייצירת התוכן הוויזואלי - visual generative capabilities



יכולת זו היא רכיב מרכזי בפתרון ה- modeSense , נועד לשפוך פרספקטיבית ויזואלית של דigramות UML .

איך זה עובד (מאתה הקלעים)?

1. כתבים תיאור טקסטואלי של מה שרצו לראות.
2. הטקסט עובר המרה לרצף של "חוקרים שונים" שנקראים tokens.
3. המודל של AI (טרנספורמר) מייצר את התמונה טוקן אחרי טוקן – ממש כמו שמצירים נקודה אחרי נקודה – אבל בצורה חכמה: כל נקודה תלואה بما שכבר צויר ובמה שנאמר בטקסט.
4. בסוף – נוצרת תמונה שלמה שמתאימה לטקסט.

איך זה מתקשר ל-UML?

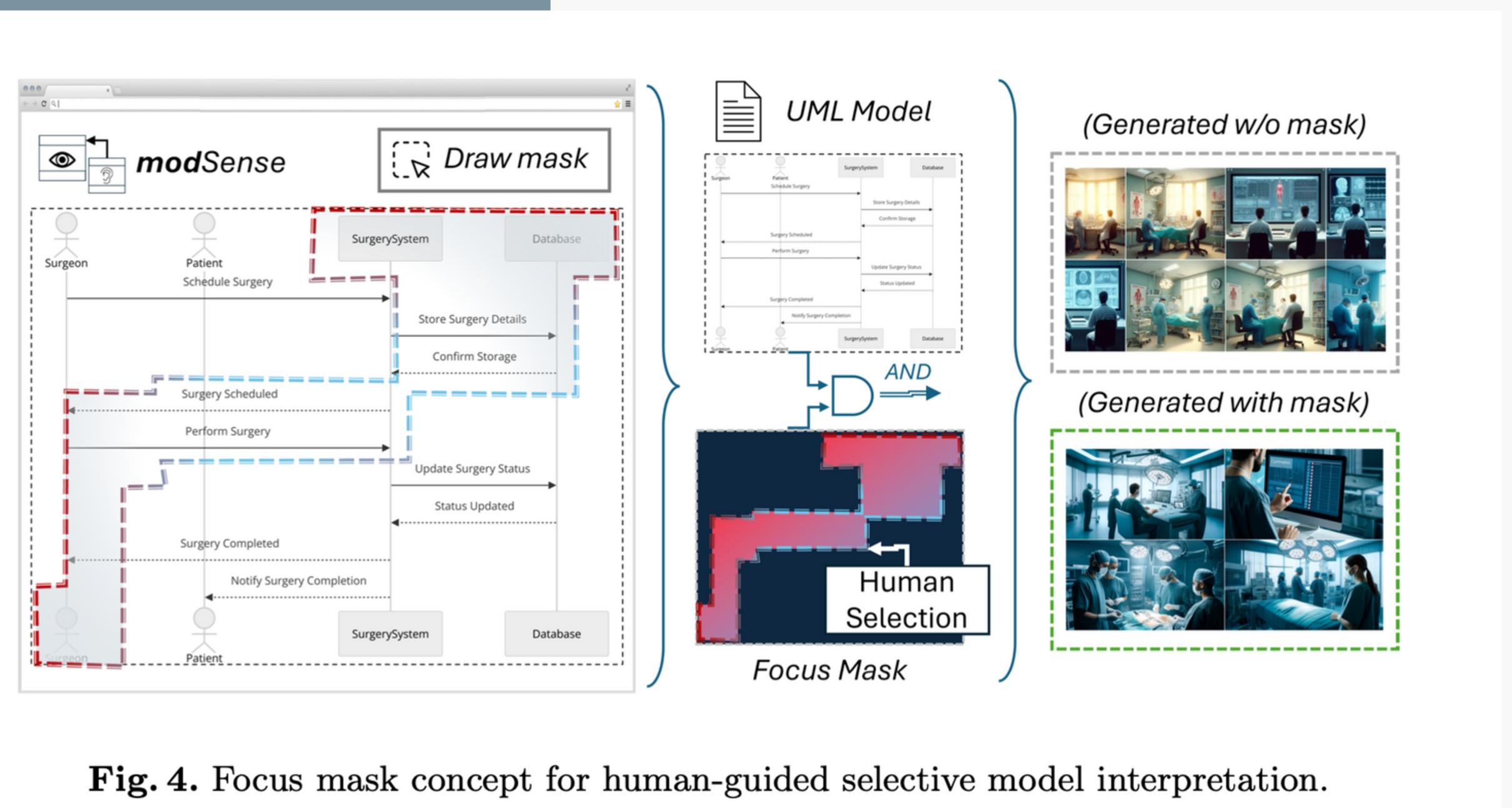
Auditory generative capabilites

Table 1. UML Key Elements for Generating *modSense* interpretations

No.	UML Diagram	Key Elements for Generating <i>modSense</i> interpretations
1	Class	Classes, Interfaces, Associations, Generalizations, Dependencies
2	Object	Objects, Links, Fields
3	Use Case	Actors, Use Cases, Relationships, System Boundary
4	Sequence	Objects, Lifelines, Messages, Activation Bars, Time Constraints
5	Collaboration	Objects, Links, Messages, Sequence Numbers
6	Activity	Activities, Transitions, Swimlanes, Forks, Joins, Start/End
7	State Machine	States, Transitions, Events, Actions, Initial and Final States
8	Component	Components, Interfaces, Ports, Connections
9	Deployment	Nodes, Artifacts, Links
10	Timing	Objects, Time Rulers, State or Value Lifelines, Events
11	Interaction	Interaction Frames, Activity Nodes, Control Flows, Decisions
12	Communication	Objects, Messages, Links, Sequence Numbers
13	Composite Struct.	Parts, Connectors, Ports, Collaborations
14	Package	Packages, Dependencies, Mergings, Access

- הדיאגרמה מומרת לתיאור טקסטואלי.
- הטקסט מקודד ל-BPE tokens – יחידות עיבוד קטנות.
- המודל הטרנספורמר בונה את הפלט שלב אחר שלב .
- נוצר קובץ שמע באמצעות מודלים כמו MusicGen או -Text-to-Speech AI.
- התוצאה: הסבר קולי מותאם לדיאגרמה, שנועד לשפר את הבנה.

focus masks



מסכת מיקוד הן תוכנה מרכזית בפתרון ה-**modSense** שנועדה לאפשר התאמת אישית עמוקה של חווית הלמידה והבנה של דיאגרמות UML.

- היא מאפשרת למשתמש להציג חלקים או אלמנטים ספציפיים בדיאגרמה לחקירה נוספת.
- לאחר שהמשתמש מגידר את מסכת המיקוד המערכת מייצרת את התוכן עבור החלק הספציפי של דיאגרמת ה-UML שהודגש על ידי המסכה.

תהליך הgenerציה זהה משתמש בטכנולוגיות AI שמסוגלות ליצור או להתאים בצורה דינמית את הייצוג הויזואלי של אותו אזור נבחר, בהתאם לקלטים של המשתמש וכללי מידול מוגדרים מראש.



— תוצאות ראשונית ממערכת *ModSense* —

מטרת הערכה:

בדיקות איות ההפקה הויזואלית והקולית של דיאגרמות UML בתחוםים מעשיים:
קידדות, קלאות ארגנית, כריכת ספרים

תוצאות חיוביות :

- ייצור תכנים ויזואליים מוצלחת לדיאגרמות פשוטות.
- הצלחה בולטת בדיאגרמות עם ממד זמן (Sequence, Activity וכו').

מוגבלות שזורה:

- המודלים שנבחרו היו פשוטים מדי, לא מיצו את יכולות המערכת.
- חוסר השקעה מצד המשתמשים בבניית דרישות איותיות לדיאגרמה.

fig.5.6.7

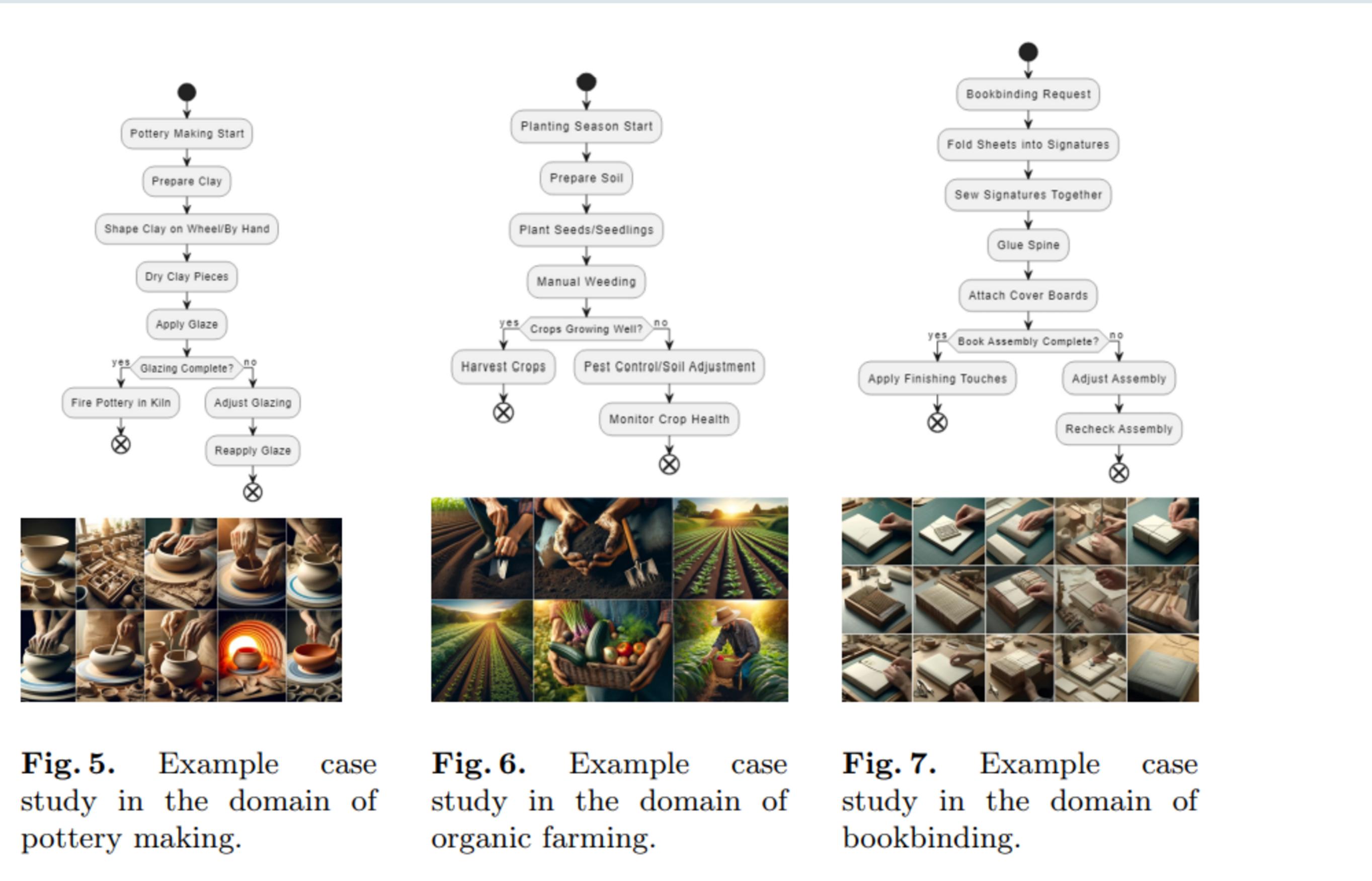


Fig. 5. Example case study in the domain of pottery making.

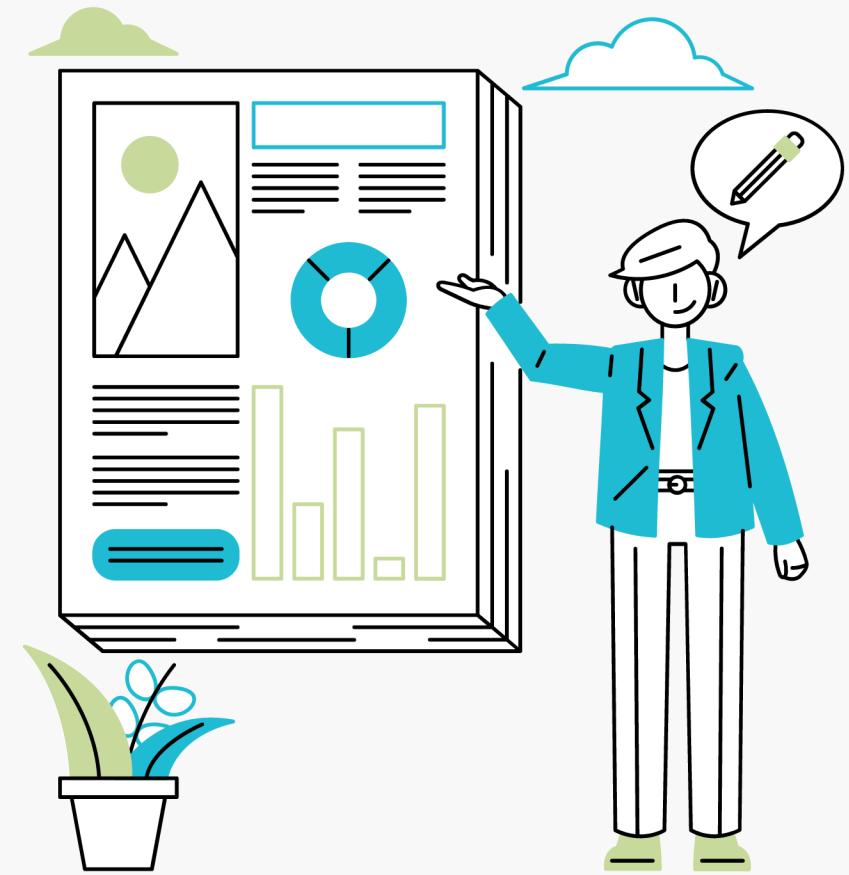
Fig. 6. Example case study in the domain of organic farming.

Fig. 7. Example case study in the domain of bookbinding.

הвисוי האמפרי – modSense4All

מטרה	אופן הביצוע:	קבוצות:	משימה:	משתתפים:	מטרת הניסוי:
המטרה של אופן הביצוע היא לספק תמונה שלמה, כמותית וaicوتית על תרומת AI למידת UML	מבחן פתיחה + (Pre-test) מבחן סיום (Post-test) מעקב אחר שימוש בחומרים (LMS) סקרים שבועיים למשוב איכוטני	קבוצת ביקורת : למידה מסורתית קבוצה ניסיונית : משאבים ויזואליים AI וקוליים מבוססי	תכנות מערכת עזר לניטוחים רפואיים שילוב בין מנוסים ב UML וכאליה שזו הפעם הראשונה שלהם	סטודנטים לתואר ראשון שני במדעי המחשב	לבוחן האם משאבי למידה שנוצרו ע"י בינה מלאכותית משפרים את הבנת UML

סיכום



המאמר מציג גישה חדשה שמשלבת טכנולוגיה מתקדמת עם צרכים חינוכיים אמיתיים. במקום ללמידה בצדקה יבשה – אנחנו מקבלים כלים שմדברים איתנו בשפה אישית, רגשית, וחזותית.
ואז זה דוגמה מעולה לשימוש חכם ב-





Thank you

