Нижче наведені ідеї та пропозиції щодо впровадження **ШІ-модуля** у вашому додатку, з урахуванням того, що це може бути **платна опція**, яку можуть купувати і СТО, і клієнти. Модуль включатиме кілька підсистем: **OCR** (розпізнавання тексту), **розпізнавання діагностичних кодів**, **голосовий інтерфейс** та **модуль початкової діагностики** на основі опису клієнта чи майстра.

**1. Архітектура та підхід до реалізації ШІ**

1. **Хмарний vs Локальний підхід**
   * Найпростіше рішення: використовувати **хмарні сервіси** (Google Cloud Vision, AWS Textract, Azure Cognitive Services, Whisper API від OpenAI тощо) для OCR, розпізнавання мовлення та ін.
   * Якщо є вимоги зберігати дані суто на власній інфраструктурі, можна інтегрувати локальні моделі (наприклад, реалізації Whisper для голосового розпізнавання або Tesseract для OCR). Але це складніше у впровадженні та масштабуванні.
2. **Модульність**
   * Кожен підфункціонал (OCR, розпізнавання голосу, декодування кодів, чатбот-діагност) краще виділити в окремий сервіс (мікросервіс, якщо потрібна масштабованість).
   * Їх можна **вмикати/вимикати** залежно від оплаченої підписки. Кожен опційний пакет — свій набір можливостей.
3. **Взаємодія з основною системою**
   * Основний додаток (ваш React/Node.js) надсилає файли або аудіо/відео-потоки у відповідний AI-сервіс.
   * Отримує результат (розпізнаний текст, перелік кодів, пропозиції щодо діагностики) і відображає користувачеві.
4. **Безпека та конфіденційність**
   * Важливо зберігати отримані дані (особливо голосові записи, скани актів) у захищеному сховищі.
   * Можливо, слід **анонімізувати** частину даних (наприклад, особисту інформацію) перед відправкою в хмарні сервіси.

**2. OCR: розпізнавання друкованих актів і висновків**

1. **Сценарій використання**
   * Клієнт або майстер робить фото/скан паперового акту чи висновку з іншого СТО.
   * Система відправляє файл у сервіс OCR. Після розпізнавання повертається текст.
   * Текст можна зберегти у внутрішній базі даних, прив’язати до замовлення чи авто, відобразити для швидкого перегляду.
2. **Технологічний стек**
   * **Google Cloud Vision** або **Tesseract.js** (відкритий інструмент) для OCR.
   * Якщо потрібне розпізнавання документів із фіксованою структурою (таблиці, поля), можна використовувати **AWS Textract** або **Azure Form Recognizer**.
3. **Мультомовність**
   * При виборі хмарних сервісів слід перевірити, чи підтримують вони потрібні європейські мови (переважно підтримують).
   * Tesseract має мовні пакети для більшості популярних європейських мов, включно з українською.
4. **Платна опція**
   * Можна реалізувати ліміт на кількість розпізнаних сторінок на місяць і пропонувати розширений пакет із більшим лімітом.

**3. Декодування діагностичних помилок (OBD-II коди та ін.)**

1. **Суть проблеми**
   * Автомобільні діагностичні пристрої (OBD-II сканери) або мобільні додатки генерують коди типу P0123, B0456 тощо. Потрібно їх розшифрувати.
   * Додатково можна зчитувати помилки з екрану смартфона, якщо користувач робить скриншот діагностичного додатка.
2. **Реалізація**
   * Зібрати **базу кодів** і їхніх описів (для різних марок авто можуть бути специфічні підкоди).
   * При надсиланні скриншота або ручному введенні коду система шукає у базі відомі помилки й видає опис (“P0420 – Низька ефективність каталітичного нейтралізатора”).
   * Розширений функціонал: пропонувати базову діагностику чи кроки вирішення (“Перевірити стан лямбда-зонда, каталізатора”).
3. **Можливе використання OCR**
   * Якщо користувач робить скриншот (або фото) екрана з кодом помилки, **OCR** може розпізнати текст і виділити ці коди.
   * Далі система шукає їх у базі.
4. **Монетизація**
   * Пропонувати базову безплатну декодування для обмеженої кількості кодів.
   * Розширений доступ (повна база кодів, додаткові підказки, рекомендації) — у платному пакеті.

**4. Голосове розпізнавання та попередня діагностика**

1. **Сценарій “Клієнт описує проблему”**
   * Клієнт включає мікрофон у додатку, каже: “При розгоні чути скрегіт зліва, особливо на 50 км/год”.
   * Система за допомогою **speech-to-text** перетворює це на текст, передає у модуль діагностики (LLM чи спеціальний рушій правил), і формує набір уточнюючих питань або можливих причин.
2. **Сценарій “Майстер озвучує інспекцію”**
   * Майстер може продиктувати: “Знайдено витік у гальмівній системі, задні колодки зношені на 80%” — система транскрибує і зберігає у картці ремонту.
   * Далі ШІ може підказати типові рекомендації (“Необхідна заміна колодок, перевірити гальмівні шланги”).
3. **Технічний стек**
   * **Whisper (OpenAI)** або **Google Cloud Speech-to-Text** для багатомовного розпізнавання.
   * **LLM-модель** (наприклад, GPT чи інші) для формування діалогових підказок і аналізу опису проблеми.
   * Локальний варіант: існують open-source реалізації Whisper, але для великих проєктів хмарні сервіси можуть бути зручніші через масштабування.
4. **Підтримка європейських мов**
   * Переконатися, що обраний сервіс (Whisper, Google Cloud) має високу якість розпізнавання потрібних мов (українська, англійська, німецька, польська тощо).
   * Можна додати **автоматичне виявлення мови** або дати користувачеві вибрати мову інтерфейсу.

**5. Інтелектуальний помічник для попередньої діагностики**

1. **Збір бази знань**
   * Чатбот (наприклад, на базі GPT) можна “підживлювати” даними про типові несправності для конкретних марок і моделей автомобілів (з урахуванням їх віку).
   * Користувач описує проблему, чатбот уточнює запитання, видає **ймовірні причини** та **рекомендації** (застереження: це не офіційна діагностика, а лише допоміжна інформація).
2. **Налаштування робочого середовища**
   * Задати обмежений контекст, щоб LLM відповідав у рамках “автомобільних консультацій”, а не розкидався зайвою інформацією.
   * Можна зберегти історію діалогів, якщо користувач дає згоду на використання даних для покращення моделей (додатково пояснити це в політиці конфіденційності).
3. **Потенційні ризики**
   * Легальне питання: якщо ШІ порадить некоректне рішення, що спричинить пошкодження авто чи небезпечну ситуацію. Має бути **дисклеймер**, що це тільки рекомендація, а фінальний діагноз ставить майстер-фахівець.

**6. Опційні пакети, ліцензування ШІ-модуля**

1. **Для клієнта**
   * **Безплатна версія**: обмежена кількість запитів до чату, базове розпізнавання 1-2 документів на місяць.
   * **Преміум**: необмежений чат для діагностики, розпізнавання актів, збереження історії, голосове введення тощо.
2. **Для СТО**
   * **Базова ліцензія**: можливість транскрибувати певну кількість аудіозаписів майстрів, OCR для внутрішніх документів.
   * **Розширена ліцензія**: автопереклад, підтримка кількох мов у діагностиці, повне інтегроване рішення з підказками для майстрів.
3. **Трекінг використання**
   * Лічильник звернень до OCR, кількості хвилин розпізнавання мовлення, звернень до LLM.
   * Пропонувати додаткові пакети, якщо клієнт перевищує ліміт.

**7. Приклад архітектурної схеми (спрощено)**

pgsql

CopyEdit

[Frontend (React)] -- User interacts

|

v

[Backend (Node.js/Express)]

| \

| \----(AI Services module, microservice or external APIs)

| 1) OCR Service (Google Cloud Vision / Tesseract)

| 2) Speech-to-Text (Whisper / Google Cloud Speech)

| 3) LLM-based Bot (OpenAI API / Local LLM)

|

v

[MongoDB]

- user data

- documents

- transcripts

- usage counters (for billing)

1. **Frontend** надсилає файли (зображення актів, аудіозаписи) чи текст.
2. **Backend** викликає відповідні AI-сервіси (вбудовані або зовнішні).
3. Результати (розпізнаний текст, рекомендації) зберігаються в MongoDB і повертаються на фронтенд.
4. **Usage counters** фіксують, скільки разів викликали OCR/стенограму/чатбот, щоб обмежувати згідно з планом підписки.

**8. Підсумок**

* **Модуль OCR**: розпізнавання друкованих актів, висновків, діагностичних скриншотів.
* **Декодування кодів несправностей**: база OBD-II + OCR для скриншотів.
* **Голосовий інтерфейс**: розпізнавання опису проблем від клієнта чи майстра, з можливістю збереження й автоматичної генерації рекомендацій.
* **Чатбот-діагностика**: LLM-асистент, який задає уточнюючі питання, пропонує потенційні причини несправності.
* **Опційні пакети**: базові функції для всіх, розширені можливості для тих, хто оформлює платну підписку (як СТО, так і кінцеві клієнти).
* **Мультимовність**: українська як основа + підтримка європейських мов (англійська, німецька, польська тощо) в OCR і розпізнаванні голосу.