

यहाँ दिए गए डेटा का हिंदी अनुवाद है:

"कनेक्टिंग द डॉट्स" चैलेंज में आपका स्वागत है 

पढ़ने को फिर से सोचें। ज्ञान को फिर से खोजें।

क्या हो अगर हर बार जब आप कोई PDF खोलें, तो वह सिर्फ़ वहीं न रहे—बल्कि आपसे बात करे, विचारों को जोड़े, और आपकी पूरी लाइब्रेरी में अर्थ बताए?

यही वह भविष्य है जिसे हम बना रहे हैं — और हम चाहते हैं कि आप इसे आकार देने में मदद करें।

"कनेक्टिंग द डॉट्स" चैलेंज में, आपका मिशन साधारण PDF को एक बुद्धिमान, इंटरैक्टिव अनुभव के रूप में फिर से कल्पना करना है—एक ऐसा अनुभव जो संरचना को समझता है, अंतर्दृष्टि को सामने लाता है, और एक भरोसेमंद शोध साथी की तरह आपको जवाब देता है।

---

## आगे की यात्रा

- राउंड 1:** चीज़ों की शुरुआत करें दिमाग बनाने से — कच्चे PDF से तेज़ गति और सटीक सटीकता के साथ संरचित रूपरेखा निकालें। फिर, इसे ऑन-डिवाइस इंटेलिजेंस के साथ शक्ति दें जो सेक्षण को समझता है और संबंधित विचारों को एक साथ जोड़ता है।
  - राउंड 2:** यह प्रदर्शन का समय है! एडोब के PDF एम्बेड API का उपयोग करके एक सुंदर, सहज पढ़ने वाला वेबएप बनाएं। आप अपने राउंड 1 के काम का उपयोग करके एक भविष्यवादी वेबएप डिज़ाइन करेंगे।
- 

## यह क्यों महत्वपूर्ण है

दस्तावेजों से भरी दुनिया में, जो जीतता है वह अधिक सामग्री नहीं है — यह **संदर्भ** है। आप सिर्फ़ उपकरण नहीं बना रहे हैं — आप हमारे पढ़ने, सीखने और जुड़ने के भविष्य का निर्माण कर रहे हैं। आपकी पृष्ठभूमि कुछ भी हो — एमएल हैकर, यूआई बिल्डर, या इनसाइट क्लिपर — यह आपका मंच है।

## क्या आप इसमें शामिल हैं?

अब लाइनों के बीच पढ़ने का समय है। डॉट्स कनेक्ट करें। और एक ऐसा PDF अनुभव बनाएं जो जादू जैसा लगे। चलिए शुरू करते हैं।

---

## राउंड 1A: अपने दस्तावेज़ को समझें

**चैलेंज थीम:** डॉक्स के माध्यम से डॉट्स को जोड़ना

**आपका मिशन** आपको एक PDF दिया जाता है – लेकिन इसे केवल पढ़ने के बजाय, आपको इसे एक मशीन की तरह समझने का काम सौंपा गया है। आपका काम दस्तावेज़ की एक संरचित रूपरेखा निकालना है – अनिवार्य रूप से **शीर्षक**, और **H1, H2, और H3** जैसे हेडिंग – एक साफ़, पदानुक्रमित प्रारूप में।

यह रूपरेखा आपकी बाकी हैकथॉन यात्रा की नींव होगी।

**यह क्यों महत्वपूर्ण है** PDF हर जगह हैं – लेकिन मशीनें स्वाभाविक रूप से उनकी संरचना को नहीं समझती हैं। एक रूपरेखा एक्सट्रैक्टर बनाकर, आप स्मार्ट दस्तावेज़ अनुभवों को सक्षम कर रहे हैं, जैसे सिमेटिक सर्च, अनुशंसा प्रणाली और अंतर्दृष्टि निर्माण।

**आपको क्या बनाने की आवश्यकता है** आपको एक ऐसा समाधान बनाना होगा जो:

- एक PDF फ़ाइल स्वीकार करे (50 पृष्ठों तक)
- निकालेः
  - शीर्षक
  - हेडिंग्स: H1, H2, H3 (स्तर और पृष्ठ संख्या के साथ)
- नीचे दिए गए प्रारूप में एक वैध JSON फ़ाइल आउटपुट करेः

JSON

```
{  
  "title": "एआई को समझना",  
  "outline": [  
    { "level": "H1", "text": "परिचय", "page": 1 },  
    { "level": "H2", "text": "एआई क्या है?", "page": 2 },  
    { "level": "H3", "text": "एआई का इतिहास", "page": 3 }  
  ]  
}
```

**आपको प्रदान किया जाएगा**

1. एक नमूना इनपुट PDF (उदाहरण, sample.pdf)
2. प्रारूप स्पष्टता के लिए एक नमूना ग्राउंड ट्रूथ आउटपुट (sample.json)
3. नमूना डॉकरफ़ाइल
4. नमूना समाधान

---

## डॉकर आवश्यकताएँ

- कृपया सुनिश्चित करें कि आपकी डॉकरफाइल **AMD64** आर्किटेक्चर के साथ संगत है। चूँकि हम एक AMD64 मशीन पर इमेज बनाएंगे और चलाएंगे, इसलिए आपकी बेस इमेज और कोई भी निर्भरता linux/amd64 का समर्थन करनी चाहिए। वैकल्पिक रूप से, आप प्लेटफ़ॉर्म को स्पष्ट रूप से निर्दिष्ट करने के लिए अपनी डॉकरफाइल में निम्नलिखित शामिल कर सकते हैं: FROM --platform=linux/amd64
  - **CPU आर्किटेक्चर:** amd64 (x86\_64)
  - कोई GPU निर्भरता नहीं
  - मॉडल का आकार (यदि उपयोग किया गया हो)  $\leq$  200MB
  - **ऑफलाइन काम करना चाहिए** – कोई नेटवर्क/इंटरनेट कॉल नहीं
- 

## अपेक्षित निष्पादन

हम निम्नलिखित कमांड का उपयोग करके डॉकर इमेज बनाएंगे:

```
docker build --platform linux/amd64 -t mysolutionname:somerandomidentifier
```

इमेज बनाने के बाद, हम प्रस्तुत निर्देशों में निर्दिष्ट रन कमांड का उपयोग करके समाधान चलाएंगे:

```
docker run --rm -v $(pwd)/input:/app/input -v $(pwd)/output:/app/output --network none mysolutionname:somerandomidentifier
```

आपके कंटेनर को चाहिए:

- /app/input डायरेक्टरी से सभी PDF को स्वचालित रूप से प्रोसेस करें, प्रत्येक filename.pdf के लिए /app/output में एक संबंधित filename.json उत्पन्न करें