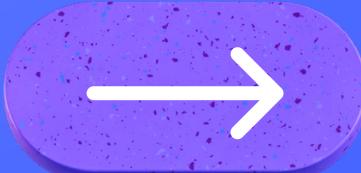


Course Crawlers

A Transformative Course
Discovery Platform for
Aggregating and Presenting
courses from Diverse Ed-tech
Platforms.



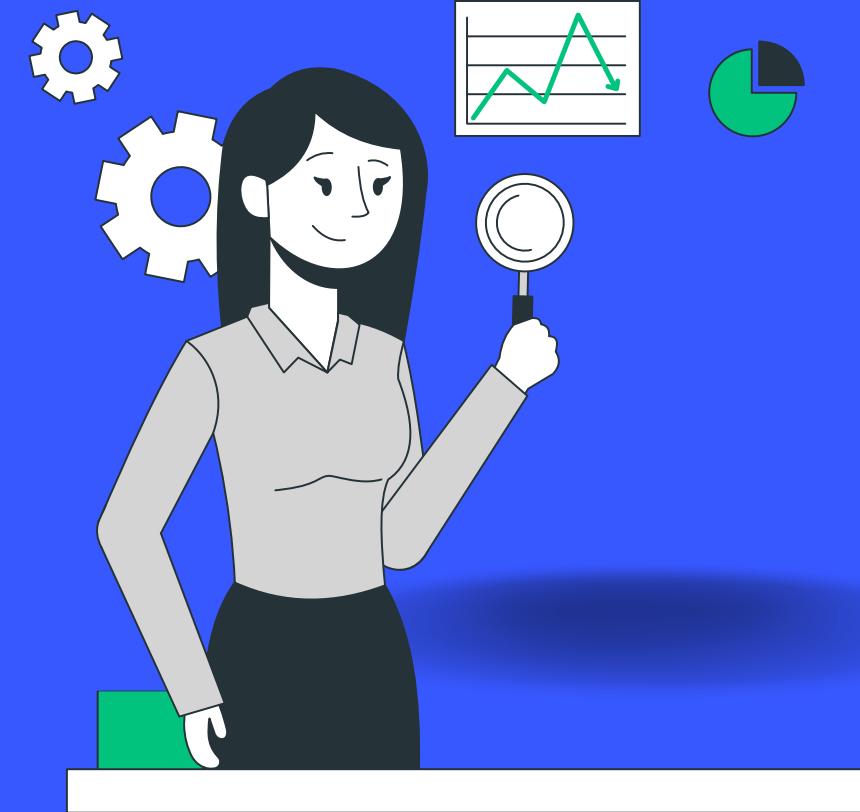
Problem Statement

In today's digital age, access to quality education is critical. Numerous online platforms offer a vast array of courses, making it challenging for learners to discover and select courses that align with their interests and goals. To address this, we aim to create a Course Discovery Platform that aggregates and presents courses from various course providers in a user-friendly manner.



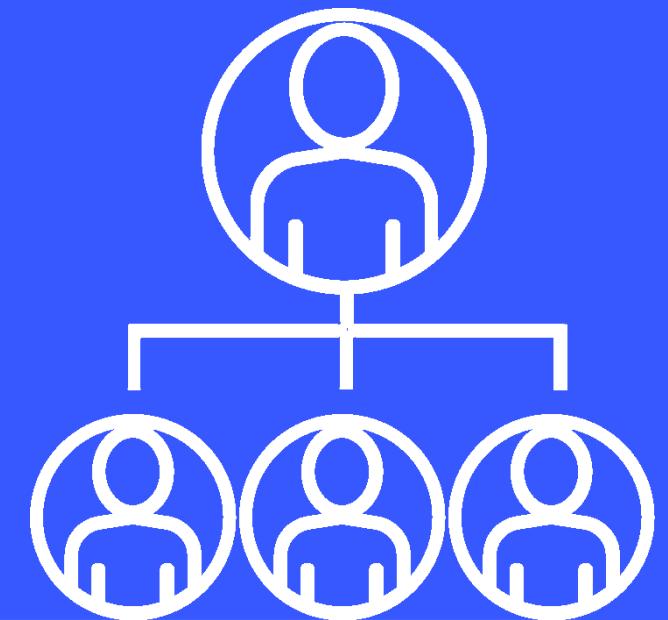
Solution

Course Crawler is an automated software or web application designed to systematically and comprehensively search, collect, and present information about courses available on various online education platforms upon users will. It offers a convenient solution for learners and educators by streamlining the process of discovering and managing courses.





Specificity



XYZ Platforms

Limited to Their Catalog: These platforms are limited to the courses available within their catalog, and learners can only choose from the courses offered directly on the platform.

Course Crawlers

Aggregators: Course crawlers are not content creators but they collect course information from various platforms , providing a centralized search and discovery service.



Need for Course Crawlers

Platform Overload

- *Too many ed- tech platforms so it gets difficult to navigate and search for courses*

Lack of Personalization

- *Difficulty in finding courses tailored to individual interests and goals*

Course Discovery Simplification

- *Users can easily search, filter, and discover courses from multiple sources, simplifying the course discovery process.*

Transparency

- *It often include user-generated reviews and ratings for courses..*





Flowchart of process

Tasks	Status
• User Input	Specify search criteria (e.g., subject, level)
• Web Scraping	Crawl course provider websites
• Data Processing	Extract course information
• Filtering	Apply user-defined filters
• Display Course Listings	Present courses to the user
• User Interaction	Click on a course for more details
• Retrieve Course Details:	Fetch additional information
• User Reviews	User Reviews: Display user-generated reviews and ratings



CC



Major Characteristics

Platform Integration

Integrate with popular course platforms like NPTEL, Coursera, edX, Udemy, LinkedIn Learning, Khan Academy, etc., to fetch a wide variety of courses.



Course Filtering

Option to filter courses based on criteria like price, duration, and difficulty level

Course Reviews and Ratings

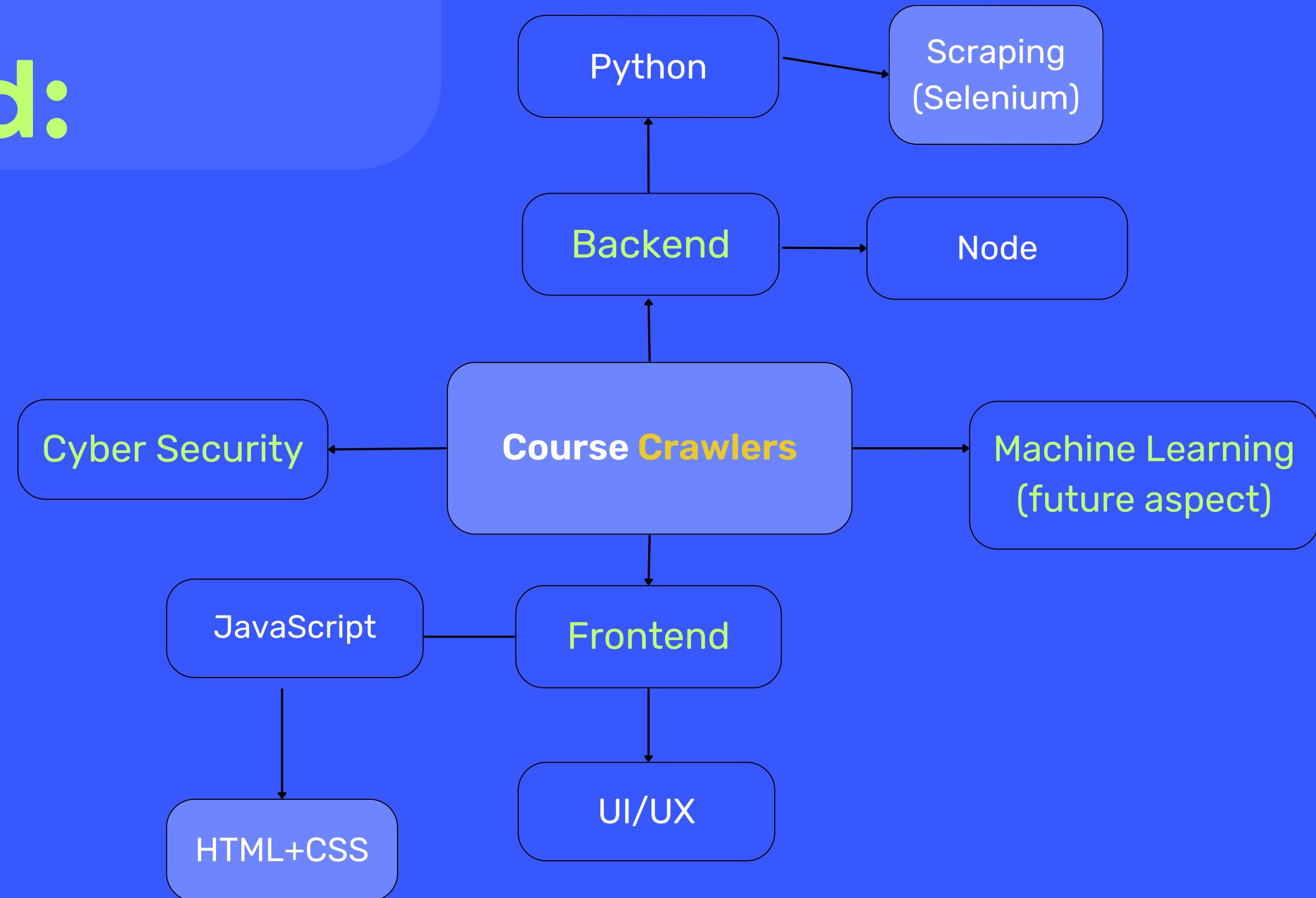
Enable users to leave reviews and ratings for courses they have taken, helping others make informed decisions.

Recommendation Engine

Implement a recommendation system that suggests courses based on user preferences, viewing history, and ratings. sent courses to the user



Tools to be Used:



UI/UX of the WebApp

Course Crawler

Search for a course:

Search

NOC:Lighter Than Air Systems Aerospace Engineering	Go To Course
NOC:Introduction To Ancient Indian Technology Aerospace Engineering	Go To Course
NOC:Introduction To Experiments In Flight Aerospace Engineering	Go To Course
NOC:Aircraft Design Aerospace Engineering	Go To Course
NOC:Computational Science In Engineering Aerospace Engineering	Go To Course
Introduction To Aircraft Control System Aerospace Engineering	Go To Course
NOC:Fundamentals Of Theoretical And Experimental Aerodynamics	Go To Course

Frontend Development Tools :

- HTML
- Tailwind CSS



RISKS

- 1. The Risk component is: WebDriver endpoint of Selenium Grid / Selenium Standalone Server. The attack vector is: Triggered by browsing to to a malicious remote web server. The WebDriver endpoint of Selenium Server (Grid) is vulnerable to DNS rebinding. This can be used to execute arbitrary code on the machine**

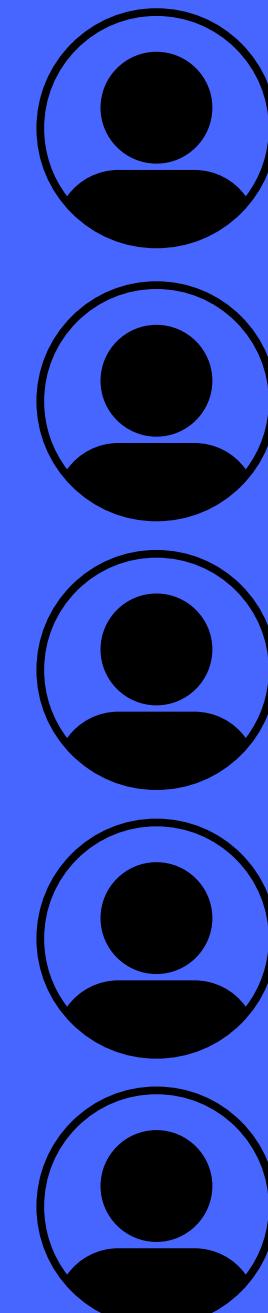
- 2. Data consistency depends on the site , which might get hinder in case if the target(course site) site is down.**

- 3. User might give fake reviews**





TEAM



MOHIT KUMAR SINGH (ML / SCRAPING)

Project Lead

ANSHUMAN SHUKLA(BACKEND)

Member

CHAITANYA SINGH KUSHWAHA(UI/UX)

Member

JHALAK JAIN (WEB SECURITY)

Member

SURYANSH DESHWAL (DATA SECURITY)

Member

HARSH JAIN (DATA SCRAPING)

Member

