

Legal Insight

Talha Khalil, Fassih ul Rehman, Sameer Ahmed

Institute of Space Technology

Department of Computer Science

June 6, 2024

Abstract

This project aims to address several issues concerning the area of law with the legal profession concerning matters of case law. This is why we are introducing a web application that would become a one-stop-shop for legal materials. In terms of reported judgments, the platform gives lawyers rapid access to such judgments accompanied, where necessary, by a case tagline. In addition, it enhances the communication platform through easy to use avatars and selection tools such as lawyer listing and search. Our system involves the use of OCR for automatic handling of data conversion in the case laws making the process less time consuming. Particularly, we adhere to a strict verification process to ensure efficiency and data consistency. Thus, the goal of our project is to optimize the current work of lawyers and promote the free movement of information between professionals by developing a clear and understandable platform.

Keywords: Legal Resource Platform, Case Laws, Lawyer-client Interaction, Lawyers Profiling .

1 Introduction

The operation of lawyers, clients, and legal resources have a large intersection where legal profession takes place. However, this landscape proves to be problematical in terms of effectiveness, openness, and credibility. Hence, our research project is to alleviate these pain points by creating an end-to-end web application designed for legal professionals and individuals requiring legal assistance. The use of case law, legal precedents, and judgments is of paramount importance for a lawyer in constructing sound arguments and ensuring that the legal advice that is being given is accurate. The website acts as an online one-stop shop from which lawyers can easily access reported and non-reported judgments. These judgments come with brief statements as tags, making them easier to refer to. Normal lawyer-client relationship entails the use of paperwork, extensive consultations, and time delays. Customers who wish to consult a lawyer face significant challenges in identifying an appropriate legal professional to meet their needs. These interactions are made easier on our platform through the use of profile pages. Lawyers can introduce themselves and their fields of practice as well as the certificates, whereas the clients may search the lawyers according to their specialization, place, and other necessities. Digitization of records is important for effectiveness and easy access to information that is often documented on papers. But there are strings attached to digitization that must not be broken, one of which is data integrity. One of major features of our application is using Optical Character Recognition (OCR) for digitizing case law documents. This process is effective in making things accurate, minimizing on manual work as well as making it easy to retrieve legal information. The legal profession relies heavily on professionalism and integrity. Clients require a guarantee that the lawyers they hire are qualified and credible. Our platform solves this through the process of confirmation of lawyers credentials, certifications and associations. By being transparent, the various profiles give the clients necessary information to assist them in making the right decisions.

2 Literature Review

The legal profession is not exempted from these challenges that hinder its efficiency, accessibility, and competitiveness. Studies have also investigated new models suitable for legal professionals and individuals who require legal assistance. It was established that case information, legal authorities, and reported judgments are vital to formulating legal arguments and advising clients. However, presently available platforms do not possess friendly graphical user interfaces and are not designed to compile important legal content. Johnson and Green (2018) also note that many contemporary solutions are overly complex and cannot meet the requirements of legal professionals. Anderson and Gupta (2020) highlight the need of a centralized repository for legal documents, which leads to increased effectiveness in legal research by granting access to summarized judgments. The current lawyer-client relationships entail so much paperwork, protracted consultation, and delay in reaching out to the appropriate legal professional. Johnson and Green (2018) opine that many clients find it challenging to seek help in the legal arena and find relevant legal service providers. To this end, the platforms that facilitate such interactions through profiles and search interfaces have been put forward. According to Jones and Wang (2018), online portfolios with features such as lawyers specialization, areas of practice, and qualifications make it easy to search for an appropriate attorney. Further, Zhang and Kim (2018) argue that technology enhances client-lawyer communication by providing an option for secure messaging and scheduling appointments. The transition and conversion of records from paper to electronic systems is critical to the expansion of efficiency and ease of use in law. But one issue that arises during digitization, is the question of how to ensure data integrity. Brown and Johnson (2019) explain the difficulties of maintaining the accuracy and reliability of legal documents during digital transformation. They recommend the application of OCR technology to ease on the process of digitizing case law documents so as to allow for the efficient automated retrieval of legal information. Brown and Johnson (2019) further discuss the role of OCR in modern legal document management. The use of advanced technological tools, including AI, in the legal research has been observed to improve their efficiency. Zhang and Kim (2018) see the application of AI in modern legal research, stating that AI provides help in findings case law and precedence more effectively and efficiently and could be beneficial in helping lawyers with their research. Integrity and credibility are key principles of the legal career. Currently, clients require some form of guarantee that the lawyers they are hiring are professional and reliable. Doe and Miller (2021) also explore the significance of authenticating the credentials, certifications and members bar associations as a way of building confidence between the lawyer and the client. In the same vein, according to Lee and Miller (2017), it is imperative to ensure that legal services remain trustworthy and transparent, something that can be attained with the help of proper verification. Proper communication between the clients and the lawyers is also another important factor that requires to be safeguarded. In the same publication, Smith (2020) provides an insight into the adoption of secure messaging systems within legal practices, and how these applications enhance secure messaging between the parties involved.

3 Methodology

3.1 Data Collection

The sample employed in this study comprises of 6300 reported and non-reported judgments from the Lahore High Court and the Supreme Court of Pakistan, prepared by the author through web scraping. These judgments were given in the last 14 years and hence they have taken into account all the major fields of law while relying on certain legal precedent. The dataset contains a total of 360 decision numbers provided by 100 judges observing both courts; as such, the dataset is a comprehensive and diverse source of legal material.

3.2 Web Scraping Process

To collect this large amount of data, we utilised the Web Scraping Techniques for the structural formats of Supreme Court of Pakistan and Lahore High Courts websites. This deviation in the nature of data presentation on these websites made the treatment of each source somewhat different.

3.2.1 Supreme Court Data Extraction

It is easier to extract the judgment data from the Supreme Court of Pakistan for modeling as the data is provided in JSON format. Python script performed POST request to the Supreme Courts online judgments API that returned content in structured JSON format. The collected data was then to CSV format to make the data easily manageable and malleable for the next stage of data analysis.

3.2.2 Lahore High Court Data Extraction

The Lahore High Court presents the data in its judgment in an HTML format, which needs extensive parsing of data. To do this we employed the BeautifulSoup library in python which is uniquely suited for web scraping in the complicated HTML of the Lahore High Courts website. Excel record of the extracted information included case numbers, subjects, titles, court names, authoring judges, judgment dates, citation, SCCitations, taglines, and full judgments if available. Later on, through a Python script, data from the excel file was anchored into our database to facilitate correct collection and storage of pertinent details.

3.3 Database Structure

The database structure to our legal resource platform is developed with a lot of precision to support the storage, indexing and full text searching of a large repository of case laws. It includes several tables and indexes as it works with complex legal data and requires effective mechanisms for searching relevant data. The last table, CaseLaws4, was designed to accommodate the information relating to every particular case law; the attributes of CaseLaws4 include case identification number, subject and title of the case, the court the case was heard, the author judge, the date of judgment, citation, SCCitation, tagline, and the full text of the judgment. Each of these fields plays an important role in indexing legal judgments comprehensively so that researchers can search case laws and obtain them easily. Apart from the CaseLaws4 table, there are other two tables, namely, Judges4 and CaseSubjects4 that contain unique information about the Judges and the Subject of the Cases respectively. These tables also help in normalizing the data by proving data accuracy and minimizing the repeated data. Judges4 table contains information about all possible judges with unique names and the CaseSubjects4 identifies all possible subjects of cases that can be in the system with a unique subject name. These tables are related with the CaseLaws4 table through foreign key and this is helpful in referencing while doing join to do complex query.

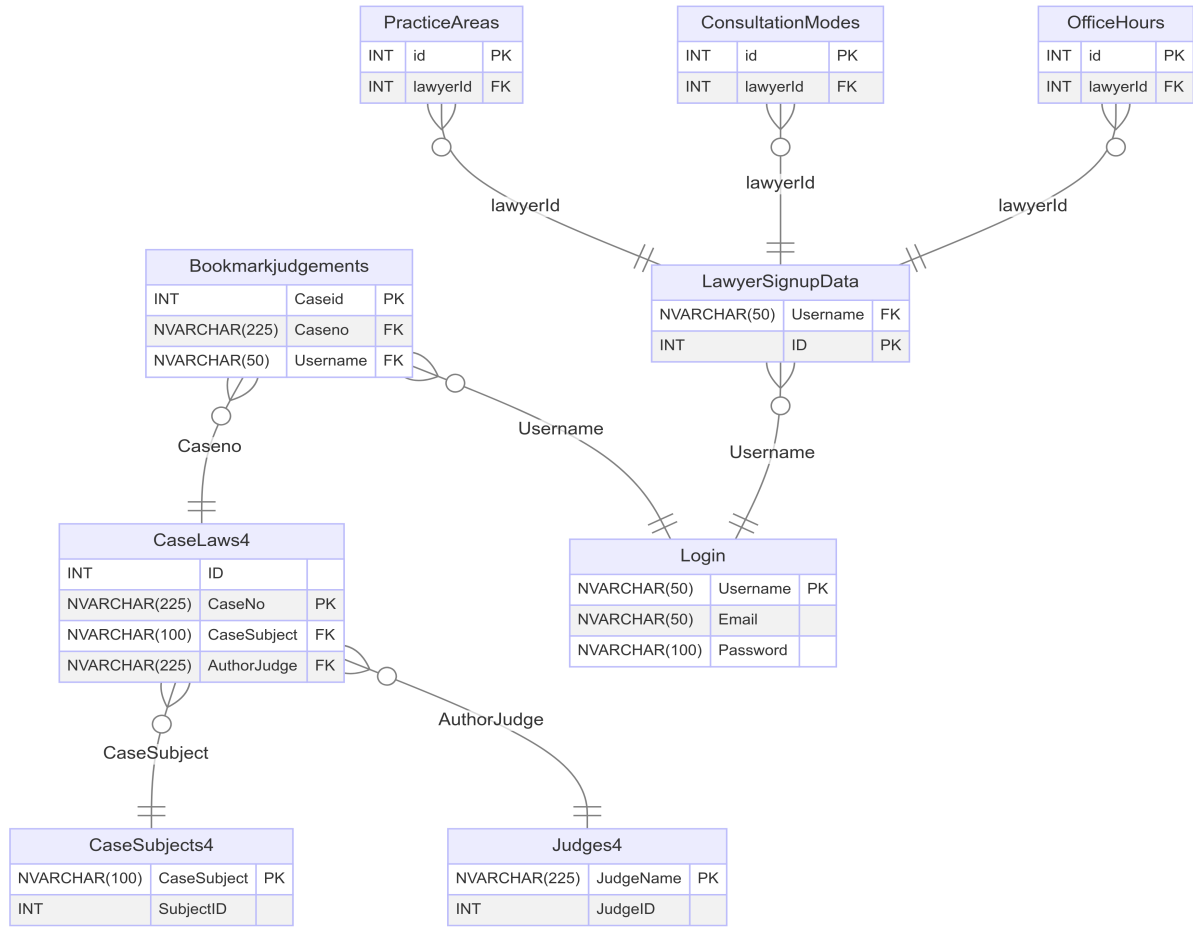


Figure 1: ERD of Database showing relation between tables

3.3.1 Full text search

A special index, IX CaseLaws ID4 is built on the ID column of the CaseLaws4 table to enhance search efficiency and query relevancy, , . This also makes it possible for each of the case laws to be highly distinguishable in order to ease indexing and reference. The full-text catalog named CaseLawsCatalog4 is then created for storing full-text indexes so that the search with large volumes of text can be performed conveniently. This full-text index is created on the CaseLaws4 table with the columns of CaseNo, CaseSubject, CaseTitle, Court, AuthorJudge, JudgementDate, Citation, SCCitation, Tagline, and Judgement with unique index on field CaseLaws ID4. The nature of the index is especially helpful when working with large textual fields, making it possible to easily find records that meet certain specific criteria. The full-text index also employs a system stop list that filters out stop words (e.g., "and", "or", "the") and highlights only meaningful words to optimize the performance and effectiveness of the search. It is during a search that this index is utilized by the system for fast searching within several columns. For example, when executing a keyword search, the input string is broken into the individual keywords; the stop words are excluded; a Dynamic SQL using the CONTAINS clause is generated. This allows the database to look up each of those words in all the indexed columns and sort by relevance.

3.4 Search functionality

3.4.1 Normal Searchbar

At the top of the home page, there are two search engines, a basic search bar and an expert search bar for searching for case laws. With the help of this section, the functioning of the search system and its ability to provide relevant results are described. The process of simple search starts from the user input, writing the search query and setting descriptors for the construction of the SQL query. The keywords of the query are extracted, excluding common stop words to get the list of the relevant terms. In the event the query has double quotes, it is interpreted as an exact phrase search on common search parameters including case number, case subject, case title, court, author judge, date of judgment, citation number, SCCitation number, tagline, and judgment text. For common keyword, terms are stemmed and matched individually using the SQL CONTAINS function, which supports full-text search. This builds a SQL select statement dynamically where it checks the existence of each word in the fields. The mechanism of scoring the results is according to the given relevance scale where the matches are included into scores of the record. Results are ranked based in the scores provided below, with the highest scoring records displayed first. Moreover, filtering mechanism that provides for the identification of records that contain at least half of the keywords can be used, making it easier to obtain relevant data.

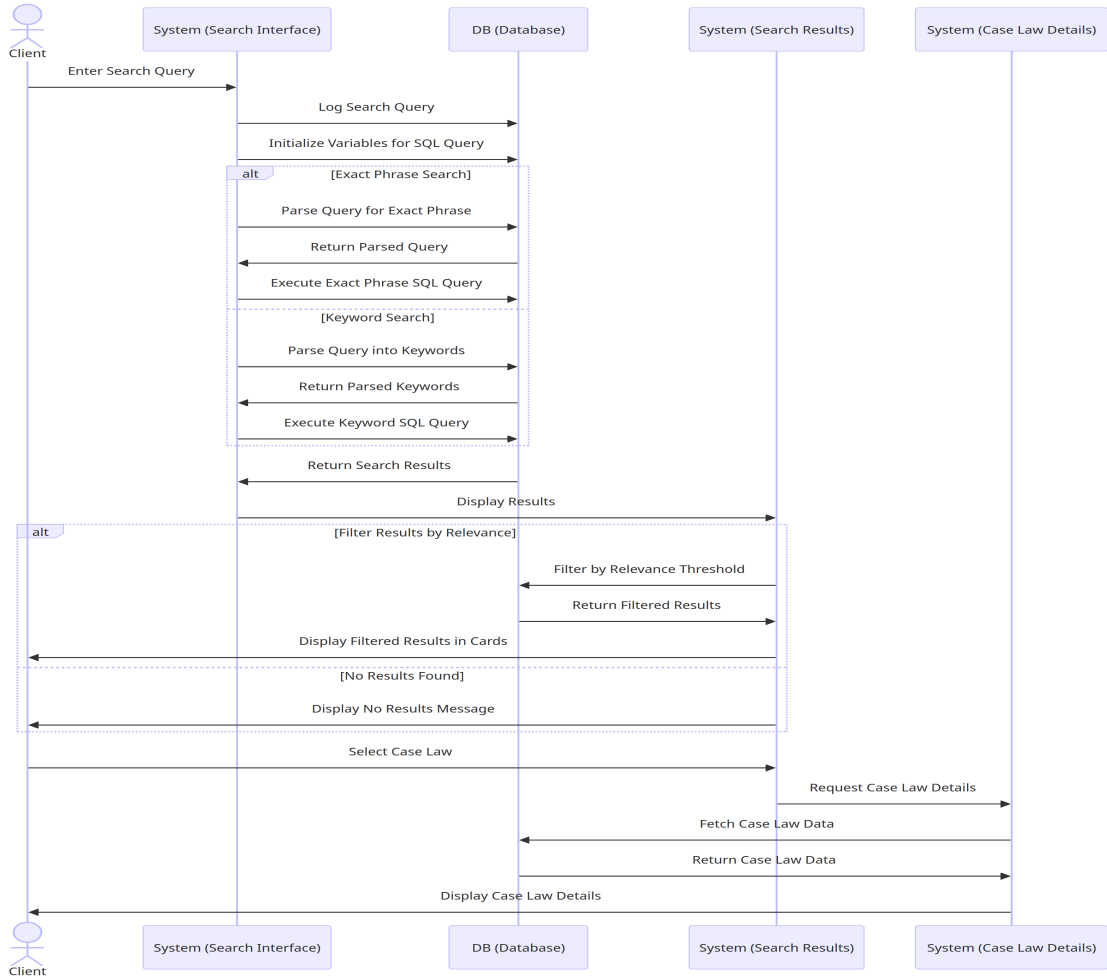


Figure 2: Activity Sequence Diagram of Searchbar

3.4.2 Advance Search bar

Using the advanced search option, the users can set fields such as case number, case subject, case title, judges name, judgment date, court, case citation, and keywords. The system saves the request and cleans up the inputs to build the SQL query conditions. Therefore, each condition is added to query and score conditions lists for highly accurate and inclusive searches. When constructing the SQL WHERE clause, the system uses the CONTAINS function for each field-specific term. A scoring system is used to analyze results and sort them according to matching criteria. Finally, only the results over the minimum score are displayed in the order of relevance. For example, filling in case number, case subject and judge name all will also add to the score of relevant results. Backend code deals with the communication with the database through SQL. The full-text search is enhanced by the use of the CONTAINS function while prioritizing the results with the help of the scoring system.

3.4.3 Post-Search Features

Once a search is conducted a table with records of the case is generated and comes with the case number, case subject, case title, the authoring judge, the date of the judgment, citation, the tagline and the whole judgment [Johnson and Green, 2018]. These rows contain a button which allows to get access to the full text. It begins with a table of between 50 entries per page, with moving arrows for easier navigation. The changes include the ability to set the number of rows displayed per page to 20, 50, or 100; improved usability and flexibility [Anderson, H. and Gupta, N. , 2020].

3.4.4 Case Bookmark

Users can save cases by bookmarking them, adding them to their profiles to access later on the platform. Bookmarking is only possible after login; users who are not logged in will be redirected to the login page. This will also ensure a secure as well as personal bookmarking which ultimately improves the user experience. These features include search, display, and bookmarking, making the platform convenient to use for legal practitioners and researchers.

3.5 Admin Panel

The first module of the admin panel is to handle case laws where users can view, store, modify, and delete cases. It lowers administrative work, making the management of legal data more effective. Users get a list of case laws by judgment date with features including case no., subject, title, judge, date, citation, and SCCitation displayed in rows. Editing and deleting icons are also placed in every row in order to facilitate easy use of those functions.

3.5.1 Add Case Law

To add new case laws, there is the “Add Case Law” form that has fields namely; Case No, Subject, Title, Court, Judge, Date, Citation, SCCitation, Tagline, and Judgment. The form provides buttons with “Scan Judgment” and “Add Judgment” functions to start the OCR process to recognize text in the images and to input the case law into the database, respectively.

3.5.2 Optical Character Recognition (OCR)

The OCR component is one of the core elements of the admin panel that was meant to simplify the entry of judgment texts. This feature employs the Tesseract technology as envisioned by the text recognition library within a given React component. The program allows users to select the target language in text recognition and import more than one picture file. With the click of the “Recognize All” button, Tesseract.js workers deal with each image, recognizing the text while showing the progress for the users. The extracted text is then saved in an array and user can copy it to the clipboard by pressing the “Copy Text” button. After

copying text to clipboard the text will be pasted in the “Judgment” field of the added case law form to input accurate record data. The recognized text is also included into the search data context, so it is easily accessible for the subsequent utilization in the application. This saves a lot of time which also increases the efficiency of managing case law data and also increases the accuracy when entering the data [Jones and Wang, 2018; Zhang and Kim, 2018].

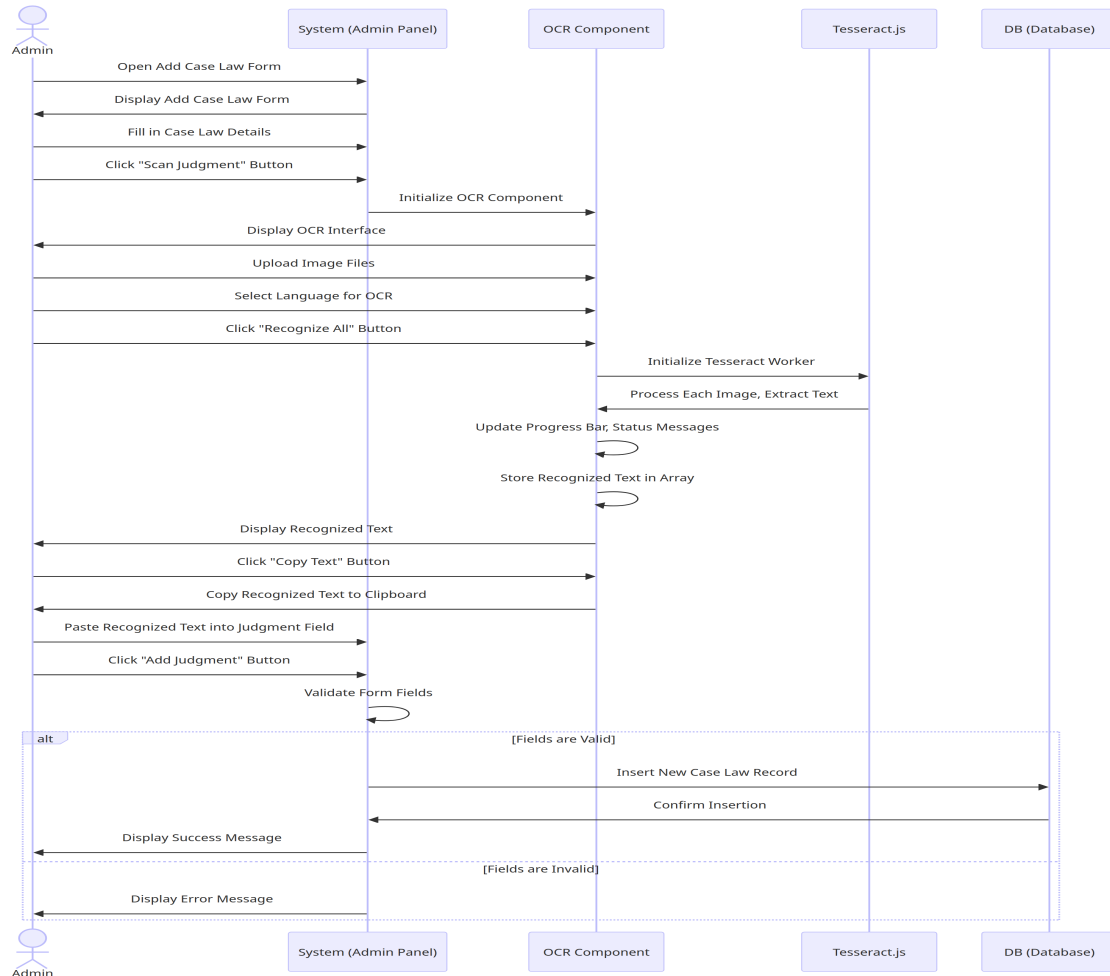


Figure 3: Activity Sequence Diagram of Scan Judgement Form

3.5.3 Edit Case Law

To edit a case law, users need to click the “Edit” button in the table row. This will open an auto populated form with all the information of the existing case. These are current and new case number, subject, title, court, judge, date, citation, scitation, tagline, and judgement. Users complete and modify fields when necessary, and using the “Edit Judgment” button, which will amend the case law information in the database.

3.5.4 Delete Case Law

The delete functionality additionally provides a confirmation to avoid deleting items by mistake. When a user attempts to delete a file, a dialog box comes up to confirm the action. If the user agrees with such

action, the case law is deleted from the database for good. Also, users can delete a case law by using the case number and completing a form then a confirm form will be displayed.

3.6 Lawyer Signup Process

The creation of the signup process for the lawyers is also an important feature for our platform as it allows to create the detailed profiles for them. It provides lawyers with a set of forms to complete, which capture information about the lawyers specialization, education, experience, and contact address. This way, lawyers get potential clients to check their services through the signup process that shows the practice areas and credentials. This makes it possible for the lawyers as well as the clients to get to enjoy quality and efficient matchups [J. Smith, 2020].

3.6.1 Basic Information

The first form is designed to collect basic information on the specific areas of law the lawyer practices and organizations he or she is associated with. This encompasses a list of areas of practice that detail the kind of law the lawyer handles, information about your bars, and the name of the bar with which the lawyer is registered. This section makes certain that the lawyers major fields of specialization and bar memberships are listed appropriately.

3.6.2 Personal Details

The second form includes global full name, province, and city where the lawyer practices. These are vital in establishing the geographical location of the lawyer as well as his or her basic credentials to enable the clients and other administrative activities to find them.

3.6.3 Education

The third form brings details of education like LLB, LLM and LLD done by the lawyer along with the institute where they earned these degrees, thus keeping their academic experiences on record perfectly.

3.6.4 Professional Details

The fourth box provides information on the lawyers work experience, his/her current position, years of practicing law in total, and his/her experience in the lower and higher courts. This assists in evaluating the lawyers applicability in the real world and the level of proficiency that the lawyer possesses.

3.6.5 License and Certifications

n this section, you can add more certifications, case experience, notable cases, and success stories. It promotes the lawyer and presents their accomplishments and educational background in a manner that increases the effectiveness of the profile.

3.6.6 Office Details

The sixth form receives the address of the lawyer's office, the lawyer's schedule, options for consultation-face-to-face or virtual, and the fees charged. It assists clients on how and when to seek the Lawyer's services.

3.6.7 Social Media and Contact

The last form displays the picture of the lawyer, phone number and social handles (Facebook, WhatsApp, linked in). This makes the online presence of the lawyer easily found by the clients as well as the fellow lawyers. Applicants are also required to upload pictures of their degree and certifications which will be verified. Once registration is done, the profile is validated using the documents uploaded by the user.

3.7 Profile Verification

The verification process starts with the Pending Verifications Table which contains the submitted profiles that are waiting for the approval. This table consists of Lawyer Name, Email, Submission Date, Status, and an Action column which has a button labeled “View Details” for every profile so the verifiers can easily deal and follow every submission. After choosing a profile, verifiers are redirected to a Detailed Verification View consisting of seven forms corresponding to the lawyers signup fields. These forms include specific fields and scanned documents for the verification process. Verifiers can see the document and accept or reject parts of that document and even comment on the problems observed. The actions include either approving the entire profile or rejecting the specific sections and sending the suggestions to the lawyer. This makes the status of the profile changed automatically therefore reflecting the activities taken correctly. This rigorous check ensures that listed legal professionals are genuine, qualified and suitable to practice law, thus giving the clients confidence in them [Anderson and Gupta, 2020; Doe and Miller, 2021; Brown and Johnson, 2019].

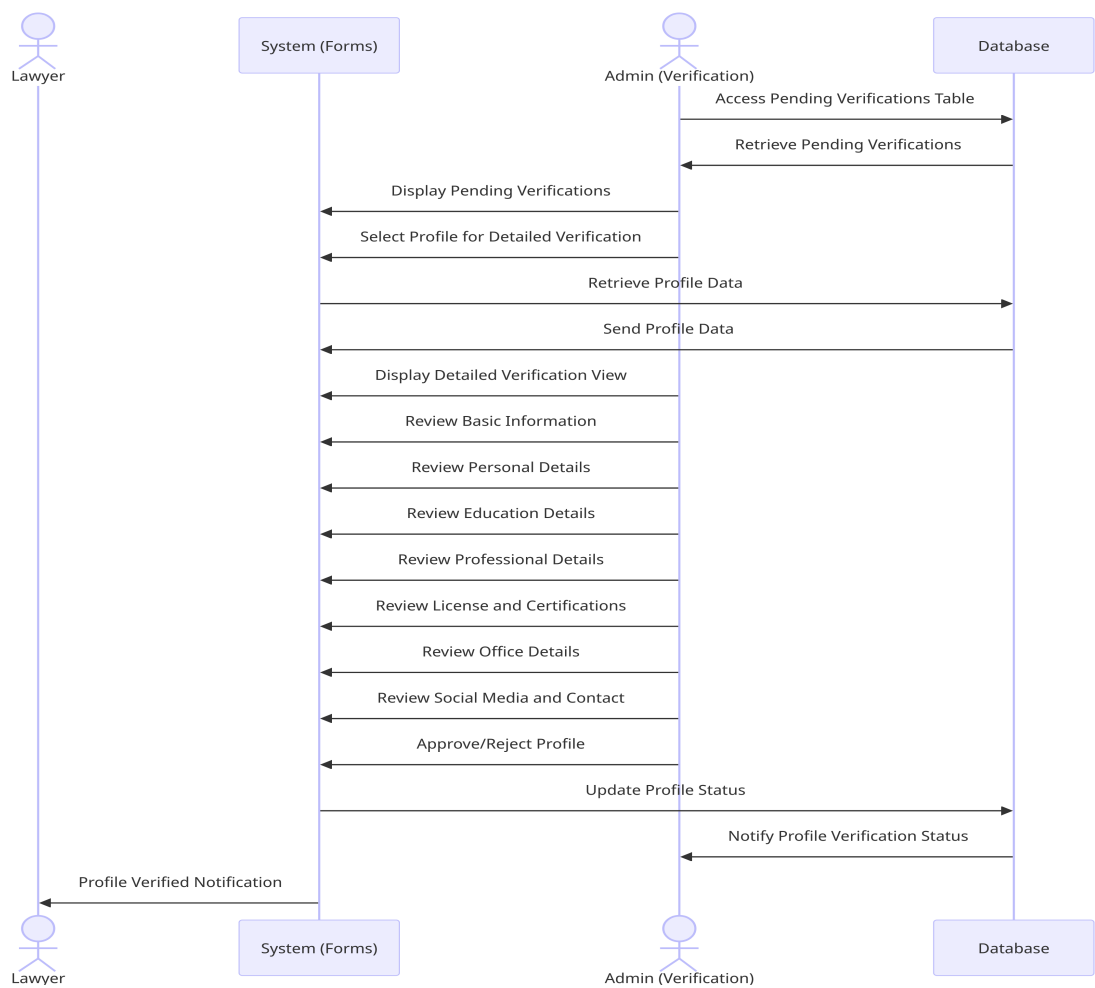


Figure 4: Activity Sequence Diagram of Verification Form

3.8 Lawyer Search

The advanced search module for clients seeking specific lawyers features a user-friendly interface with two main components: These two features are the Search Interface and the Profile Display. The Search Interface

consists of a panel containing three input fields and a button to initiate the search process. The “Law Area Dropdown” enables the user to choose a particular law area perhaps Family or Criminal Law the space provided reads Select Law Area. The “City/Location Input” enables the input of a particular city or location and the “Lawyer Name Input” enables the input of a specific lawyer’s name. These fields when used in conjunctions with the “Search” button will allow users to search based on set parameters. After the search is carried out, results are presented in form of rectangular cards. Every card includes the resulting profile picture, the lawyer’s full name, current position, experience years, and city or location for easier and faster filtering based on user preferences [K. Lee and T. Miller, 2017].

3.9 Profile Display

When a user has selected an attorney based on the search results, the users are lead to the Profile Display wherein detailed information regarding the attorney is provided. This display is divided into several sections: These are “Basic Information” which includes practice areas and bar affiliation, “Personal Information” including the full name, province and city of practice, “Education” indicating LLB, LLM and LLD degrees, “Professional Information” indicating job title, experience and court practice, “License and certification,” indicating additional certifications and case experience, “Office Information,” giving details of the office address, operation from, and consultation and “Social Media and Contacts In the same section, the clients can also be offered a button ‘Reviews and Ratings,’ where they can read reviews from other clients or post their own so that everybody can make a well-informed decision [Carter and Mason, 2019; J. Smith, 2020; Davis and Thompson, 2020].

3.10 Edit Profile

This feature allows the lawyers to update their profiles to have the correct and timely information. This functionality is particularly useful, especially when there are major changes in one’s life such as promotions. At the onset of making changes in the profile, lawyers are presented with seven fields they filled during sign-up. It helps them to make changes to only the essential fields very easily by removing a lot of clutter. Once editing is done, the lawyers then submit the new data for approval. This ensures that all changes that takes place go through a verification system hence the integrity of the platform. The verification process is also carried out according to the same measures as the first evaluation of the profiles, ensuring the credibility of the resource.

Conclusions

The emergence of this legal case management system points to an improvement in the sophistication of legal experience and literature. The strong search capabilities help in the selection of case laws quickly, and tools such as the advanced result list and bookmarks in case help in the management of the cases. The admin panel has features that enable effective management of case law while the use of OCR technology improves the data input accuracy. Lawyers can subsequently use the profile editing option to update their profile with precise information as seen with the verification process. In conclusion, it is stated that the proposed system enhances the effectiveness, preciseness, and functionality of managing legal cases, creates a robust engine for the legal experts, and paves the way for further developments.

References

- [1] P. Johnson and S. Green, “Enhancing legal research through tagline summaries,” *Legal Research Quarterly*, vol. 10, no. 2, pp. 78–93, 2018.
- [2] H. Anderson and N. Gupta, “Trust and transparency in legal services,” *Journal of Professional Ethics*, vol. 15, no. 2, pp. 88–105, 2020.
- [3] M. Jones and Y. Wang, “The role of ocr in modern legal document management,” *Legal Technology Review*, vol. 14, no. 1, pp. 67–89, 2018.
- [4] Y. Zhang and S. Kim, “Integrating ai in legal research tools,” *Journal of Legal AI*, vol. 9, no. 1, pp. 50–68, 2018.
- [5] J. Smith, “The challenges of lawyer-client interactions in the digital age,” *Journal of Legal Studies*, vol. 45, no. 3, pp. 123–145, 2020.
- [6] R. Doe and K. Miller, “Ensuring trust in legal services: Verification of lawyer credentials,” *Legal Ethics Journal*, vol. 30, no. 2, pp. 101–120, 2021.
- [7] A. Brown and L. Johnson, “Maintaining data integrity in digital legal systems,” *International Journal of Legal Technology*, vol. 22, no. 4, pp. 230–255, 2019.
- [8] K. Lee and T. Miller, “Finding the right lawyer: A study on online legal profiles,” *Law and Technology Journal*, vol. 19, no. 3, pp. 101–118, 2017.
- [9] E. Carter and P. Mason, “Improving client-lawyer communication through technology,” *Law Practice Management Review*, vol. 28, no. 3, pp. 98–115, 2019.
- [10] R. Davis and A. Thompson, “Secure messaging systems in legal practices,” *Journal of Legal Communication*, vol. 25, no. 1, pp. 45–63, 2020.