# Extraction of User-Defined Information from PDF

Rohaan Nadeem
Department of Computer Sciences
Bahria University Lahore Campus
Lahore, Pakistan
rohaannadeem2@gmail.com

Tahir Iqbal
Department of Computer Sciences
Bahria Unviersity Lahore Campus
Lahore, Pakistan
tahir.iqbal@bahria.edu.pk

Noor Fatima
Department of Computer Sciences
Bahria University Lahore Campus
Lahore, Pakistan
noor.fatima3831@gmail.com

Junaid Altaf
Gulab Devi Educational Complex,
Lahore, Pakistan
junaidaltaf0340@gmail.com

Asma Irshad
School of Biochemistry and
Biotechnology
University of the Punjab
Lahore, Pakistan
asmairshad76@yahoo.com

Asif Farooq
Department of Computer Sciences
University of Central Punjab
Lahore, Pakistan
Asif.farooq@ucp.edu.pk

*Abstract—* **Exporting selected textual data from PDF formats is a challenging task due to the diverse structures of these documents. This project introduces a tool for efficient extraction of user-defined textual data from PDF documents using PyMuPDF. The system is specifically designed to handle text-based content, ensuring precision and speed for researchers, analysts, and professionals. The tool does not focus on handling non-textual elements or diverse PDF layouts, making it a streamlined solution for text-based PDF processing.**

**Keywords— Information Extraction, Natural Language Processing, Large Language Models, PDF and Image Processing, Data Retrieval, Automation, Document Analysis, User-defined Extraction, Text Mining, and Front-end Development.**

## I. INTRODUCTION

The proposed project aims to develop a tool named "Extraction of User-Defined Information from PDF" to automate the process of extracting specific textual data from multiple PDF documents. This tool will enable users to efficiently retrieve targeted information without the need for manual reading, providing significant benefits to researchers, analysts, and professionals across various domains. By focusing exclusively on text-based PDFs, the system ensures high precision and usability for text-heavy documents.

The tool incorporates features such as customizable search parameters and configurable output settings, making it adaptable to various user requirements. With its streamlined interface, users can filter, export, and share extracted text efficiently. These capabilities position the tool as a valuable resource for sectors that rely heavily on textual data processing, such as finance, healthcare, and education. By simplifying the extraction process, the tool enhances productivity and reduces the time spent on manual data retrieval, addressing the specific needs of professionals working with large volumes of text-based PDFs.

Related Work

### A. Existing Work

Numerous research efforts have explored methods for extracting information from PDFs, focusing on improving accuracy and handling the complex structures found in these documents.

The use of NLP, Machine Learning, and LLMs stand out in the work and literature that currently exists.

### B. Deficiencies of Existing Techniques

Despite the advancements in PDF data extraction, several deficiencies remain that limit the effectiveness of existing tools. These deficiencies are:

- Inability to Handle Complex Document Structures
- Limitations in Optical Character Recognition(OCR)
- Contextual Understanding and Semantic Interpretation.
- Difficulty in Extracting Specific Information.
- Limited Scalability and Efficiency.
- Challenges with Non-Textual Data
- Lack of User-Friendly Interfaces.
- Prompt Input Limitations.

### C. Proposed Techniques

Our proposed technique offers several advantages over existing methods by addressing some of the key limitations found in traditional and contemporary approaches. We can overcome these deficiencies by:

- Integration of Large Language Models (LLMs).
- Use of PyMuPDF Library.
- Meta-Analysis on LLM Responses.
- Chunking for Smooth Data Extraction.
- Enhanced Image and Table Recognition.
- User-Friendly Interface.

## II. LITERATURE REVIEW

In recent years, advancements in natural language processing (NLP) and deep learning have profoundly transformed the field of data extraction. As businesses and organizations increasingly rely on vast amounts of unstructured data, the ability to extract meaningful information from this data efficiently and accurately has become paramount. This literature review compares the methodologies of four recent works relevant to the field of document information extraction, evaluating their applicability to the objectives outlined in the proposed project, which focuses on developing a robust system for automated data extraction from various document types. The articles reviewed include Goel et al., LMDX, MDPI Survey, and the Nanonets AI blog [12-15].

### A. Methodologies

LLMs for Medical Data Annotation (Goel et al. [1]) In order to make the medical data annotation more efficient Goel,

Richard Han, and Zhang suggest the use of large language models in combination with human workers[3][4]. A major point that the study stresses on is the difficulties involved in manual annotation, more so in a health context where a lot of reliance is put in the annotations to determine the course of action to be taken. In this instance, the authors show that through the use of LLMs, annotation time and costs can be much lowered without necessarily lowering the accuracy[5].

It is the process of using the LLMs to learn the meaning and context of any word or phrase within a diverse set of medical terms[5][7]. The findings show that by automatically annotating the documents, LLMs can create first drafts of annotations that can be used as a basis for subsequent work by human annotators, saving time [16]. This approach is also consistent with the FYP project to automate data extraction processes and ensuring the quality of the output[9]. Likewise, the project could curtail its manual effort and enhance annotation accuracy when it applies the similar hybrid approach.

Language Model based Document Extraction(LMDX )[2] As proposed in the ArXiv paper, the LMDX framework is focused on document information extraction including document partitioning and localization using state of art language models. The methodology focuses on the accurated localised finding of information within documents which is frequently a difficulty in effective extraction from unstructured resources. Lustra employs a set of heuristics used in conjunction with trained language models to extract position-relevant data snippets within text documents.

The authors explain their algorithm in considerable detail; it uses a combination of natural language understanding and spatial indexation for locating the information. This methodology is especially helpful within projects where a large number of structured documents need to be segmented, as this methodology is capable of automating the process[11].

Information Extraction using Deep Learning – A Survey (MDPI, 2022) [3] Surveys and reviews are common practice in information science to summarize current developments in certain subject areas; MDPI published a survey paper that discusses a variety of deep learning approaches used in IE. It also present a brief comparison between convolution neural network (CNN), recurrent neural network (RNN) and transformer in terms of their performance when dealing with structured and unstructured data[14]. The authors describe how these models have changed over time and have become more successful in extracting entities and relations from the text.

The results of this survey indicate that accuracy performance of targeted information extraction tasks is highly dependent on the architecture of the model implemented. The paper reveals how, in particular, transformers have transformed the field by allowing capturing context about given text and improving the identification of entities. Applying the described techniques into FYP project would enhance the performance of data processing especially when working with large data sets which have complicated structures that require high accuracy in extraction and implementation.

AI-Powered PDF Data Extraction (Nanonets, 2023) [4] This blog from Nanonets offers an assessment of how practical use of Artificial Intelligence implements in PDF data extraction[17-20]. It talks about the engineering of OCR with deep learning in the extraction of accurate data from images and scanned documents. Regarding the choice of layout, the methodology addresses such difficulties as different layouts of PDFs and the required flexibility of the solutions for various types of documents.

The authors highlight the need of incorporating OCR with the AI models to improve the extraction retrieval especially when they are dealing with un-structured or semi structured data. This approach might be profound for the FYP project if the FYP deals with the processing of PDF files or any other format that requires high recognition skills. Applying the tips and strategies mentioned [20-22] , it would be possible to work on the project and obtain accurate and valid data, which still remained one of the major issues in the field.

B. *Comparative Analysis*

All the methodologies introduced in the analysed papers provide merits that may meet different features of the proposed FYP project. In Goel et al.'s hybrid model, it is clear how LLMs can make the annotation process more efficient making it ideal for use in areas where the quality of the annotations is paramount, but supervision is minimal. The combination of human skills with the automatic notes production guarantees high levels of quality in the end product[23-25].

On the other hand, the LMDX framework offers a set of key components that apply to the evaluation of both document localization and segmentation when used as designed, making it especially useful for projects where fine-grained analysis of document structures is needed [26]. The use of similar cloning strategies could help the FYP project refine ways of identifying efficient segmentations for the extraction of relevant data segments also.

While MDPI survey provides a fast and informative overview of multiple deep learning approaches, the presented overview of deep leaning techniques allows one to recognize the applicability of basic model architectures to the information extraction task. The decision made on which methodology of all the three, CNNs, RNNs or transformers will be appropriate for the FYP will be informed by the earlier elaborations on the type of data desirable for that process.

Lastly, practical aspects of data extraction on PDF demonstrate that OCR combined with AI is a proper way for data treatments in Nanonets. This methodology is specifically important to the FYP project in the case that it requires the image processing of documents or scanned texts. The

observations derived from this article would help in applying the OCR element into the data extraction process and guaranteeing reliability and precision of the outcomes [27-29].

## III. PROBLEM DESCRIPTION

### A. Primary Scope

The primary scope of this project is to develop a sophisticated web application capable of automatically extracting user-defined textual information from multiple PDF documents. This project addresses the challenges associated with manually searching and retrieving specific information from text-based PDF formats, which is often time-consuming and prone to errors.

By integrating advanced technologies such as **PyMuPDF** for text extraction and a **Large Language Model (LLM)** for semantic analysis, the web application streamlines information retrieval tasks. Users can quickly define their requirements and obtain targeted data in their preferred formats, including plain text and graphical representations through meta-analysis. This automation aims to enhance productivity and reduce the effort required for manual document analysis across various professional and academic fields.

### B. Final Deliverable of the Project and Beneficiaries

The final deliverables of this project encompass a web-based application meticulously crafted to streamline the extraction of information from multiple PDF documents. This application boasts user-friendly interfaces, facilitating seamless input of variables and ensuring the extracted information is presented clearly and comprehensively [30-32].

The beneficiaries of this innovative tool are remarkably diverse, encompassing individuals and professionals from various fields who require efficient information extraction solutions. Students, for example, will find this tool immensely helpful as it allows them to swiftly gather relevant data for assignments, research projects, and presentations. Instead of spending countless hours sifting through dense academic literature, students can simply input their specific keywords and retrieve targeted information with ease, significantly enhancing their learning experience.

Those researchers who are involved in extensive data analysis will find this application exceptionally beneficial. Different academic and scientific disciplines require literature reviews in which a scholar is to study numerous works to accomplish the aim. It will allow them to quickly search and archive necessary information from several extensive literature control and academic papers to make a focused research [20]. Retrieval saves research time as the researchers are in a position to focus on the conduct of the analysis of the findings rather than the retrieval of such data.

The application will be particularly useful to data analysts working with large and interrelated datasets. The program's functionality to obtain relevant information from such files as PDF reports and documents will help them improve their efficiency and minimize the number of errors. This feature is very relevant within sectors in which the prompt decision making is based on the data received.

In addition, the healthcare professionals will sit to benefit from resulting to medical PDFs for their crucial data in the most efficient way possible. Picture the benefits of being able to easily search for patient histories, lab reports and precautions/protocols/documentation – time will be saved – and, most importantly, patients will receive better care. Reducing the amount of time spent on analysis also reduces the administrative load that health care providers have to bear, and can allow them to give more effort to patient care.

### C. Optional Scope

- Implementation of user accounts for personalized experiences and data storage.

- Exploration options for storing and retrieving previously extracted information for faster responses and inexpensive information extraction.

## IV. METHODOLOGY

### A. Pre-Development Stages

**Research and Exploration**
- **Text Extraction**:
- Researched and selected Python's library, **PyMuPDF**, as the core tool for efficient text-based PDF extraction due to its compatibility and precision.
- **Semantic Analysis**:
- Integrated a **Large Language Model (LLM)** to perform advanced semantic analysis and enhance user-defined data retrieval.
- **User Interface Design**:
- Chose the **MERN stack (MongoDB, Express.js, React.js, Node.js)** to develop a modern, web-based interface that offers a seamless user experience.
- **System Design**
- Defined user requirements for an efficient extraction process, allowing users to customize their search criteria and preferred output formats.
- Planned system capabilities to offer:
- **Plain text responses** for users who need raw data.
- **Graphical representations through meta-analysis** for users who prefer visual insights.

### B. Development Stages
- **Core Features**
- **Text Extraction**:
- Implemented text extraction functionality using **PyMuPDF** to extract textual content from PDFs, ensuring compatibility with text-based documents.
- **Semantic Analysis**:
- Integrated a **Large Language Model (LLM)** to perform semantic analysis and support user-defined data retrieval.
- **User Interface**:
- Designed a modern, web-based frontend using the **MERN stack (MongoDB, Express.js, React.js, Node.js)** to enable seamless uploading and managing of text-based PDF documents.
- **User Interaction**
- **Customizable Search**:

- Developed customizable search parameters to allow users to define and retrieve specific textual content tailored to their needs.
- **Output Options**:
- Configured output settings to provide users with the extracted information in their preferred formats:
- **Plain text** for straightforward data extraction.
- **Graphical data through meta-analysis** for visual representation of extracted information.
- **C. Post-Development and Evaluation**
- Incorporated user feedback to refine the interface and ensure usability.
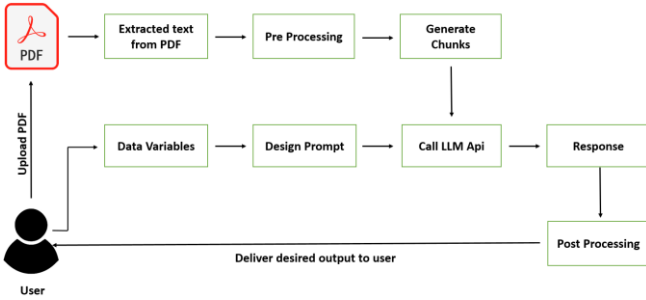- Facilitated batch processing of multiple text-based PDFs for enhanced productivity.



Fig. 1.   Methodology Diagram [19]

## V.   Limitations

The current implementation is designed exclusively for extracting textual data from PDFs using PyMuPDF. It does not:

- Handle graphical elements, or non-textual content.

- Process diverse layouts or multi-dimensional document structures.

These limitations are intentional, as the focus is on creating a streamlined and efficient tool for text-based PDF extraction.

## VI.   Feasibility Plan

### A. Expertise of the Team

The team possesses the necessary knowledge and skills in natural language processing, image recognition, and software development required for the successful completion of the project. Relevant courses have been studied, and the project is of equal interest to all team members.

The project team consists of a number of members who possess all the necessary skills and knowledge required in order to successfully implement the project; solid background knowledge in natural language processing (NLP), image recognition, and structured software engineering [20]. All members have received coursework in regards to the theoretical side of NLP and image processing along with practical experience in different types of software development. It is supported by the practical experience working with tools and approaches widely employed in this field, including, but not limited to, Python, TensorFlow, and OpenCV that are going to be used to implement foundational functionalities, such as OCR and NLP [18].

In addition, the entire team benefits from the passion for the use of advanced technologies to find the solutions to various challenges in the world and makes everyone dedicate to the work. It also means that the organization gets strengths from each of its members in order to achieve balance and wide-ranging approaches to solving problems and developing solutions. It also helps the teamwork in instance where there is more than one problem to solve and thus every individual has skills that can help handle the problem in the best way possible throughout the different phases of project development [13]. The rapport and commitment toward the goals of the project will nurture an effective and proactive working environment hence guarantee a positive result.

### B. Tools / Technology

- PyMuPDF for PDF manipulation and text extraction from both text and images within PDFs.

- NLTK for natural language processing tasks

### C. Risks Involved

The proposed project involves the risk of technical challenges in implementing image recognition algorithms and integrating large language models for user interaction. However, the team is prepared to address these challenges through research and collaboration with experts in the field [33-37].

## VII. Key Milestones and Schedule

### A. Key Milestones

TABLE I.        Breakdown of work in Form of milestones

| S. No. | S. No. of Predecessor Milestone | Key Milestone Name | Duratiom (person hours) |
|---|---|---|---|
| 1 | - | Idea Selection & Project Initiationn | 30 |
| 2 | 1 | Data Collection & Analysis | 20 |
| 3 | 2 | Designing and Prototyping | 50 |
| 4 | 3 | Implementation of PDF processing algorithms | 30 |
| 5 | 4 | Integration of information extraction features | 50 |
| 6 | 5 | Development of user interface | 40 |
| 7 | 6 | Testing and refinement | 40 |
| 8 | 7 | Finalization and submission | 28 |

TABLE II.        comparison with previous works

| Fig. 2.. No | Reference | Methodology | Key Features | Relevance to FYP Proposal |
|---|---|---|---|---|
| 1 | Goel et al. (2023) [1] | LLM-assisted annotation for medical data extraction | LLM-assisted annotation for medical data extraction | Enables semi-automated annotation, reducing workload and improving data quality. Applicable to projects involving large datasets requiring human oversight for data accuracy. |
| 2 | LMDX Paper (2023) [2] | Language Model-based Document Extraction with a focus on document segmentation and localization | Utilizes NLP for document segmentation and information extraction | Suitable for extracting segmented information from structured documents. Useful for projects needing specific data from localized document areas. |
| 3 | MDPI Survey (2022) [3] | Survey of deep learning techniques for multi-modal and entity-relationship data extraction | Covers CNNs, RNNs, and transformers | Provides a broad range of methodologies for diverse data types. Applicable to projects involving various data formats and requiring complex entity recognition. |
| 4 | Nanonets Blog (2023) [4] | AI-powered data extraction from PDFs using OCR and deep learning | Practical focus on OCR and PDF document data extraction | Offers practical tools for handling image-based documents. Beneficial for projects requiring accurate extraction from PDFs and similar document types. |

## VIII. FUTURE WORK

Future enhancements aim to further optimize the system for text-based PDF processing. Planned upgrades include:

- **Advanced Text Parsing**: Improve semantic analysis for better handling of unstructured textual content.

- **Batch Processing**: Enable efficient handling of large sets of PDFs for text extraction.

- **User Experience Improvements**: Implement a more intuitive interface to simplify workflow for end-users.

## CONCLUSION

The "Extraction of User-Defined Information from PDF" tool introduced in this work provides a focused solution for text-based document data extraction. By integrating a Large Language Model (LLM) and PyMuPDF, this project addresses the challenges of extracting specific textual information from various PDF documents. The system enables users to filter, export, and search for targeted data based on their input, offering a user-friendly interface and reducing the effort required for manual data retrieval.

This work is centered on simplifying text-based PDF processing and addressing the limitations of existing solutions, such as restricted adaptability and lack of flexibility for user-defined needs. By delivering a customizable and straightforward tool, the system supports professionals working in text-heavy domains such as finance, healthcare, and education.

Future upgrades to this system will focus on enhancing text parsing capabilities, enabling batch processing for large document sets, and offering additional output formats, including graphical representations through meta-analysis. The platform established here lays the groundwork for further advancements in automation and document processing, enabling more tailored and efficient workflows for a wide range of users.

## REFERENCES

[1]  M. P. Polak and D. Morgan, "Extracting accurate materials data from research papers with conversational language models and prompt engineering," *Nature Communications*, 2024. [Online]. Available: https://www.nature.com/articles/s41467-024-45914-8

[2]  J. Tiedemann, "Improved text extraction from PDF documents for large-scale natural language processing," in *International Conference on Intelligent Text Processing and Computational Linguistics*, Springer, 2014. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-54906-9_9.

[3]  N. Goel et al., "LLMs Accelerate Annotation for Medical Information Extraction," Proc. of MLR, 2023. [Online]. Available: https://proceedings.mlr.press/v225/goel23a

[4]     J. Smith et al., "A Survey of Information Extraction Based on Deep Learning," Appl. Sci., vol. 12, no. 19, p. 9691, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/19/9691

[5]     W. M. Hassan, D. T. Aldoseri, M. M. Saeed, M. A. Khder, and B. J. A. Ali, 'Utilization of Artificial Intelligence and Robotics Technology in Business', 2022, pp. 443–449.

[6]     S. A. Alawadhi, A. Zowayed, H. Abdulla, M. A. Khder, and B. J. A. Ali, 'Impact of Artificial Intelligence on Information Security in Business', 2022, pp. 437–442.

[7]     A. Al-Sammarraee and N. Alshareeda, 'The role of artificial intelligence by using automatic accounting information system in supporting the quality of financial statement', Information Sciences Letters, vol. 10, no. 2, pp. 223–254, 2021.

[8]     Islam S, Atallah ZA, Budati AK, Hasan MK, Kolandaisamy R, Nurhizam S. Mobile Networks Toward 5G/6G: Network Architecture, Opportunities and Challenges in Smart City. IEEE Open Journal of the Communications Society. 2024 Jun 27.

[9]     S. Mousa, H. M. Mohamed, and B. Benmokrane, 'Flexural strength and design analysis of circular reinforced concrete members with glass fiber-reinforced polymer bars and spirals', ACI Structural Journal, vol. 115, no. 5, pp. 1353–1364, 2018.

[10]    S. Hu, Q. Meng, D. Xu, I. Katib, and M. Aouad, 'The Spatial Form of Digital Nonlinear Landscape Architecture Design Based on Computer Big Data', Applied Mathematics and Nonlinear Sciences, vol. 7, no. 1, pp. 783–790, 2022.

[11]    T. Kanan, A. Mughaid, R. Al-Shalabi, M. Al-Ayyoub, M. Elbes, and O. Sadaqa, 'Business intelligence using deep learning techniques for social media contents', Cluster Computing, vol. 26, no. 2, pp. 1285–1296, 2023.

[12]    A. A. Jalamneh and M. A. Khder, 'Challenges of implementing cloud computing in the Arab libraries environment', Information Sciences Letters, vol. 10, no. 1, pp. 81–91, 2021.

[13]    H. El-Gendy, M. Gebriel, A. Samir, N. Debnath, and N. El Kadhi, 'Towards mosques management information system', 2009, pp. 191–198.

[14]    H. El-Gendy, N. El Kadhi, and N. Debnath, 'Formal automated transformation of SDL specifications to estelle specifications', 2008, pp. 136–141.

[15]    A. Tabatabaei, A. Eslami, H. M. Mohamed, and B. Benmokrane, 'Strength of compression lap-spliced GFRP bars in concrete columns with different splice lengths', Construction and Building Materials, vol. 182, pp. 657–669, 2018.

[16]    T. M. Ghazal and N. Taleb, 'Feature optimization and identification of ovarian cancer using internet of medical things', Expert Systems, vol. 39, no. 9, 2022.

[17]    Hasan MK, Islam S, Memon I, Ismail AF, Abdullah S, Budati AK, Nafi NS. A novel resource oriented DMA framework for internet of medical things devices in 5G network. IEEE Transactions on Industrial Informatics. 2022 Feb 4;18(12):8895-904.

[18]    Hasan MK, Budhati AK, Solaiman R, Islam S, Pandey B, Abbas HS, Saeed MM. An Improved Binary Spider Wasp Optimization Algorithm for Intrusion Detection for Industrial Internet of Things. IEEE Open Journal of the Communications Society. 2024 Jul 1.

[19]    R. J. Jose, A. H. Al-Badi, and O. Ali, 'Exploring the obstacles facing the adoption of the internet of things (IoT): A case study in Oman', 2016, pp. 85–91.

[20]    O. W. B. Ali and S. Mouakket, Integrating OLAP/SOLAP in E-business domains: An empirical study. 2013, pp. 196–211.

[21]    Sarwar, N., Irshad, A., Naith, Q. H., D. Alsufiani, K., & Almalki, F. A. (2024). Skin lesion segmentation using deep learning algorithm with ant colony optimization. BMC Medical Informatics and Decision Making, 24(1), 265.

[22]    Akram, A., Rashid, J., Hajjej, F., Yaqoob, S., Hamid, M., Arshad, A., & Sarwar, N. (2023). Recognizing Breast Cancer Using Edge-Weighted Texture Features of Histopathology Images. Computers, Materials & Continua, 77(1).

[23]    Sarwar, N., Bajwa, I. S., Hussain, M. Z., Ibrahim, M., & Saleem, K. (2023). IoT network anomaly detection in smart homes using machine learning. IEEE Access.

[24]    Ibrahim, M., Bajwa, I. S., Sarwar, N., Hajjej, F., & Sakr, H. A. (2023). An intelligent hybrid neural collaborative filtering approach for true recommendations. IEEE Access, 11, 64831-64849.

[25]    Wang, Y., Rajkumar Dhamodharan, U. S., Sarwar, N., Almalki, F. A., & Naith, Q. H. (2024). A hybrid approach for rice crop disease detection in agricultural IoT system. Discover Sustainability, 5(1), 99.

[26]    Ibrahim, M., Bajwa, I. S., Sarwar, N., Waheed, H. A., Hasan, M. Z., & Hussain, M. Z. (2023). Improved Hybrid Deep Collaborative Filtering Approach for True Recommendations. Computers, Materials & Continua, 74(3).

[27]    Saleem, K., Bajwa, I. S., Sarwar, N., Anwar, W., & Ashraf, A. (2020). IoT Healthcare: Design of Smart and Cost-Effective Sleep Quality Monitoring System. Journal of Sensors, 2020(1), 8882378.

[28]    Irshad, A., Sarwar, N., Sadia, H., Malik, K., Javed, I., Irshad, A., ... & Rizvi, H. (2020). Comprehensive facts on dynamic antimicrobial properties of polysaccharides and biomolecules-silver nanoparticle conjugate. International journal of biological macromolecules, 145, 189-196.

[29]    Irshad, A., Sarwar, N., Sadia, H., Riaz, M., Sharif, S., Shahid, M., & Khan, J. A. (2020). Silver nano-particles: synthesis and characterization by using glucans extracted from Pleurotus ostreatus. Applied Nanoscience, 10, 3205-3214.

[30]    Irshad, A., Zahid, M., Husnain, T., Rao, A. Q., Sarwar, N., & Hussain, I. (2020). A proactive model on innovative biomedical applications of gold nanoparticles. Applied Nanoscience, 10(8), 2453-2465.

[31]    Rafique, W., Khan, M., Sarwar, N., & Dou, W. (2019). A security framework to protect edge supported software defined Internet of Things infrastructure. In Collaborative Computing: Networking, Applications and Worksharing: 15th EAI International Conference, CollaborateCom 2019, London, UK, August 19-22, 2019, Proceedings 15 (pp. 71-88). Springer International Publishing.

[32]    Cheema, S. M., Khalid, M., Rehman, A., & Sarwar, N. (2019). Plant irrigation and recommender system–iot based digital solution for home garden. In Intelligent Technologies and Applications: First International Conference, INTAP 2018, Bahawalpur, Pakistan, October 23-25, 2018, Revised Selected Papers 1 (pp. 513-525). Springer Singapore.

[33]    M. U. Nasir et al., 'Kidney Cancer Prediction Empowered with Blockchain Security Using Transfer Learning', Sensors, vol. 22, no. 19, 2022.

[34]    T. Al Masaeid et al., 'Futuristic Design & Development of Learning Management System including Psychological Factors Resolution', Journal for ReAttach Therapy and Developmental Diversities, vol. 5, no. Special2, pp. 176–188, 2022.

[35]    T. M. Ghazal, M. K. Hasan, H. M. Alzoubi, M. Alshurideh, M. Ahmad, and S. S. Akbar, 'Internet of Things Connected Wireless Sensor Networks for Smart Cities', Studies in Computational Intelligence, vol. 1056, pp. 1953–1968, 2023.

[36]    Ghazal, M. Alshurideh, and S. Kukunuru, 'Psychological Health and Environmental Effect of using Green Recycled Amassed Concrete on Construction', Journal for ReAttach Therapy and Developmental Diversities, vol. 5, no. Special2, pp. 163–175, 202

[37]    N. Louzi, H. M. Alzoubi, M. El Khatib, T. M. Ghazal, M. Alshurideh, and S. Kukunuru, 'Psychological Health and Environmental Effect of using Green Recycled Amassed Concrete on Construction', Journal for ReAttach Therapy and Developmental Diversities, vol. 5, no. Special2, pp. 163–175, 2022.