

manage their documents efficiently and effectively. Leveraging advanced machine learning and NLP techniques, the system provides users with the ability to search, redact, and process large volumes of documents quickly and accurately. The system's three tracks, Discovery, Compliance, and Documents, cater to different document management needs, providing users with a versatile and flexible system. Additionally, the system's web-based interface and integration with AWS services such as S3 and Lambda make it easy for organizations to manage their documents at scale, without the need for costly on-premise infrastructure. Overall, the Cloud based Document Understanding System offers organizations a powerful set of tools for managing their document workflows, improving efficiency, and ensuring compliance with data protection regulations.

References

- [1] W. Li, S. Neullens, M. Breier, M. Bosling, T. Pretz, and D. Merhof 2014 Text recognition for information retrieval in images of printed circuit boards. IECON 2014 - 40th Annual Conference of the IEEE Industrial Electronics Society.
- [2] A. Muliantara, A. Sanjaya N., M. Widiarth I., and A. Setiawan I. M. 2015 Prototype of cloud-based document management for scientific work validation. International Conference on Information & Communication Technology and Systems (ICTS).
- [3] Singh D., Saini J. P., and Chauhan D. S. 2015 Hindi character recognition using RBF neural network and directional group feature extraction technique. International Conference on Cognitive and Information Processing (CCIP). Computing
- [4] J. Pradeep, E. Srinivasan, and S. Himavathi 2012 Performance analysis of hybrid feature extraction technique for recognizing English handwritten characters. World Congress on Information and Communication Technologies.
- [5] D. Nasien, H. Haron, and S. Yuhaniz S. 2010 Support Vector Machine (SVM) for English Handwritten Character Recognition. Second International Conference on Computer Engineering and Applications.
- [6] B. Liu, X Yu., P. Zhang, A. Yu, Q. Fu, and X. Wei 2018 Supervised Deep Feature Extraction for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(4), 1909–1921.
- [7] R. Ramanathan, S. Ponmathavan, V. L. Thaneshwaran N., Nair A. S., and P. Soman K. 2009 Optical Character Recognition for English and Tamil Using Support Vector Machines. International Conference on Advances in Computing, Control, and Telecommunication
- [8] Zhang et al. (2020). "Cloud Document Management Systems: A Review of Recent Advances and Challenges".
- [9] Wang et al. (2022). "Cloud-Based Document Understanding: A Comprehensive Review and Analysis".
- [10] Lee et al. (2021). "Cloud-Based Document Understanding: Recent Advances and Future Directions".