



# Adoption of Intelligent Automation Tools in the Banking Industry: A Case Study on Leveraging Indico for Document Digitization and Data Extraction

Karthika Gopalakrishnan

Data Scientist  
karthika.gopalakrishnan@cgi.com

---

## ABSTRACT

This paper investigates the adoption of intelligent automation tools in the banking industry, with a focus on the application of Indico in digitizing critical documents and extracting data elements. Indico, a cutting-edge platform powered by BERT-based models, offers advanced capabilities for document processing and automation. We explore how this tool can be effectively utilized for digitizing signature cards, W8, W9, legal documents, death certificates, and extracting data from diverse sources such as invoice documents, financial documents, rent rolls, and correspondences. Additionally, the paper discusses the accelerated adoption of automation through transfer learning mechanisms, showcasing the potential for faster industry-wide implementation.

**Key words:** Indico, Document Digitization, Data Extraction, Intelligent Automation, Financial services

---

## 1. INTRODUCTION

In an era marked by the relentless evolution of technology, the banking sector stands at the forefront of transformative change, navigating the delicate balance between tradition and innovation. One of the pivotal shifts propelling this industry forward is the increasing adoption of intelligent automation tools. As financial institutions grapple with the ever-expanding volumes of data and the intricate nature of document processing, the imperative for efficiency, accuracy, and compliance has never been more pronounced.

The banking landscape, once characterized by manual and time-consuming processes, is now witnessing a paradigm shift towards intelligent automation. This transition is driven by the compelling need to streamline operations, enhance precision in document handling, and ensure unwavering compliance with a labyrinth of regulatory requirements. The escalating complexity of financial transactions, coupled with the demand for expeditious decision-making, underscores the urgency for technological solutions capable of augmenting traditional workflows.

Efficiency is not merely a convenience but a strategic imperative in the contemporary banking environment. Intelligent automation tools, powered by cutting-edge technologies such as natural language processing and machine learning, offer a panacea for the industry's efficiency challenges. Rapid advancements in these tools empower financial institutions to automate routine, time-intensive tasks, freeing up human resources for more intricate decision-making processes and customer-centric initiatives.

Moreover, the demand for accuracy in document handling and data extraction is not only a matter of operational efficacy but a critical element in safeguarding the integrity of financial processes. The potential repercussions of inaccuracies in financial documentation are far-reaching, affecting customer trust, regulatory compliance, and overall financial stability. Intelligent automation tools, with their ability to meticulously process vast datasets,

mitigate the risks associated with human errors, ensuring a heightened level of precision in the intricate world of banking transactions.

Compliance, an enduring cornerstone of the banking sector, has become even more formidable in the face of evolving regulatory landscapes. The intricacies of compliance demand an adaptive and dynamic approach, one that intelligent automation tools are uniquely poised to deliver. By seamlessly integrating with existing systems and continuously learning from regulatory updates, these tools provide a resilient mechanism for banks to stay ahead of compliance requirements, fostering a culture of accountability and transparency.

In essence, the growing significance of intelligent automation tools in the banking sector is an inherent response to the multifaceted challenges that financial institutions face. This paper explores the role of these tools, with a specific focus on Indico, in meeting the imperatives of efficiency, accuracy, and compliance in document handling and data extraction.

As the banking industry charts its course into a digital future, the integration of intelligent automation emerges as a cornerstone, promising not only operational optimization but a paradigm shift in how financial processes are conceived and executed. In this paper, we discuss about how intelligent automation platform like Indico are adopted in the banking industry.

In this paper, we discuss about how intelligent automation platform like Indico are adopted in the banking industry.

## 2. EVOLUTION OF AUTOMATION TOOLS IN BANKING

The evolution of automation tools within the banking sector reflects a compelling narrative of technological progression, responding to the ever-growing demands for efficiency, accuracy, and compliance. Traditional banking operations were historically characterized by manual, paper-based processes that often led to delays, errors, and increased operational costs. Recognizing the inefficiencies inherent in these methods, the industry gradually embraced automation to streamline operations and enhance overall performance.

Early automation tools in banking primarily focused on routine and transactional processes, such as account reconciliation and simple data entry. These tools, although groundbreaking in their time, were limited in their capacity to handle the complexity and diversity of financial documents that modern banking entails. As the volume and intricacy of financial transactions burgeoned, a new wave of automation tools emerged, driven by advancements in artificial intelligence and machine learning.

In recent years, the banking industry has witnessed a surge in the adoption of intelligent automation tools that harness the power of natural language processing, computer vision, and sophisticated algorithms. These tools transcend the capabilities of their predecessors, offering not just automation but intelligent decision-making capabilities. Document digitization and data extraction, once manual and error-prone processes, have become pivotal arenas for the application of these advanced automation technologies

### 2.1 Challenges

Despite the progress made in automation tools, the banking industry faces persistent challenges in effectively digitizing documents and extracting crucial data. The diversity of document types, formats, and languages presents a formidable obstacle. Legal documents, signature cards, tax forms (W8, W9), death certificates, and various financial records exhibit significant variability, making it challenging to design a one-size-fits-all solution.

Moreover, the need for compliance with regulatory standards further complicates the landscape. Frequent changes in regulations and the introduction of new compliance requirements necessitate an agile and adaptive approach to automation. Legacy systems, still prevalent in many financial institutions, pose an additional hurdle, as integrating new tools seamlessly with existing infrastructure becomes a complex task.

Inaccuracies and inconsistencies in data extraction from documents can have far-reaching consequences, impacting decision-making processes, customer relationships, and regulatory adherence. The banking industry, therefore, grapples with the dual challenge of achieving high precision in data extraction while simultaneously ensuring the scalability of these processes to handle vast datasets.

### 2.2 How does Indico overcome these challenges?

Indico emerges as a pivotal player in addressing the multifaceted challenges associated with document digitization and data extraction in the banking sector. Leveraging state-of-the-art BERT-based models, Indico

excels in understanding the context, semantics, and intricacies of diverse documents, transcending the limitations of conventional automation tools.

Indico's architecture, with its robust natural language processing capabilities, is adept at handling the intricacies of legal documents, signature cards, tax forms, and various financial records. Its ability to adapt to different document types and structures ensures a versatile and scalable solution for document digitization, offering the banking industry a tool that can evolve with changing document landscapes and compliance requirements.

The following sections of this paper will delve into specific applications of Indico, exploring its efficacy in digitizing various document types and extracting critical data elements. By doing so, we aim to provide a comprehensive understanding of how Indico contributes to overcoming the challenges inherent in document processing within the dynamic realm of modern banking.

### 3. INDICO DOCUMENT DIGITIZATION

#### 3.1 Overview of Indico's Capabilities: BERT-based Models for Document Processing

Indico, at the forefront of intelligent automation, stands as a beacon in the banking industry's pursuit of enhanced document processing and data extraction. Central to its prowess are the BERT-based models, a sophisticated framework grounded in the transformative capabilities of Bidirectional Encoder Representations from Transformers (BERT). Indico's commitment to innovation manifests in the following key capabilities:

**Bidirectional Understanding:** Indico's BERT-based models leverage bidirectional processing, enabling a contextual understanding of language. Unlike traditional models that analyze text in a unidirectional manner, BERT considers the entire context of a word or phrase, enhancing its ability to comprehend nuanced meanings and relationships within a document.

**Contextual Semantic Understanding:** The semantic richness achieved by Indico's BERT-based models allows for a deeper comprehension of the context in which words and phrases appear. This contextual understanding is particularly crucial in the banking sector, where documents often contain intricate language, legal terminology, and financial jargon that demand precise interpretation.

**Multilingual Competence:** In a globalized banking environment, multilingual competence is imperative. Indico's BERT-based models are designed to effectively process documents in various languages, ensuring adaptability to the linguistic diversity prevalent in international financial transactions. This capability positions Indico as a versatile solution for institutions with a global reach.

**Adaptive Learning:** Indico's models are characterized by adaptive learning mechanisms, enabling them to continuously evolve and improve with exposure to new datasets. This adaptability is pivotal in addressing the dynamic nature of banking documents, which may undergo frequent updates, modifications, or changes in format and structure.

**Optimized for Document Processing:** While BERT-based models have demonstrated their prowess in various natural language processing tasks, Indico optimizes these models specifically for document processing within the banking sector. This targeted optimization ensures that the models excel in accurately extracting information from complex financial documents, including legal contracts, tax forms, and other crucial records.

**High Accuracy and Precision:** The application of BERT-based models enhances the accuracy and precision of Indico's document processing capabilities. By capturing intricate contextual relationships within documents, the models significantly reduce the likelihood of errors, providing financial institutions with a reliable tool for document digitization and data extraction.

#### 3.2 Use cases

Indico's BERT-based models, tailored for document processing within the banking sector, offer a versatile solution for digitizing a range of critical documents. The following examples showcase the applicability of Indico in transforming manual document handling processes into efficient, automated workflows

##### 3.2.1 Signature Cards

**Challenge:** Signature cards, integral to banking operations, often come in diverse formats with handwritten signatures, making digitization complex.

**Solution:** Indico's BERT-based models excel in recognizing and interpreting handwritten signatures. Through contextual understanding and semantic analysis, Indico accurately extracts and digitizes signature card information, ensuring a seamless transition from paper-based to digital records. This enables banks to maintain a secure and organized database of customer signatures for various transactions.

### 3.2.2 W8 & W9 forms

**Challenge:** Tax-related forms such as W8 and W9 contain intricate details and require precision in data extraction for regulatory compliance.

**Solution:** Indico's models, adept at contextual understanding, navigate the complexities of tax forms effortlessly. The BERT-based architecture captures the nuanced language, ensuring accurate extraction of relevant information such as taxpayer details, tax classifications, and associated data. This capability streamlines tax reporting processes, enhancing compliance and reducing the risk of errors.

### 3.2.3 Legal Documents

**Challenge:** Legal documents, diverse in structure and language, demand a nuanced approach for accurate digitization and data extraction.

**Solution:** Indico's contextual understanding proves invaluable in deciphering legal jargon and complex language structures. By leveraging its BERT-based models, Indico extracts key information from legal documents, including clauses, obligations, and contractual details. This enables banks to digitize legal agreements, ensuring quick retrieval and analysis for compliance and legal scrutiny.

### 3.2.4 Death Certificates

**Challenge:** Death certificates contain sensitive information and often vary in format, posing challenges in standardizing the digitization process

**Solution:** Indico's adaptability to different document structures allows for effective extraction of pertinent details from death certificates. The BERT-based models discern context, ensuring accurate digitization of information such as deceased person's details, cause of death, and relevant timestamps. This capability aids banks in managing deceased customer records with precision and compliance.

These illustrative examples demonstrate how Indico's capabilities extend beyond conventional automation tools, addressing the unique challenges posed by critical banking documents. The contextual, semantic understanding facilitated by BERT-based models ensures that Indico adapts to the intricacies of each document type, fostering a reliable and efficient digitization process within the banking sector.

## 4. INDICO DATA EXTRACTION

Indico's robust BERT-based models provide a comprehensive solution for extracting data elements from diverse sources, ensuring accuracy and efficiency in the banking sector's document processing workflows. The following section delves into specific applications, illustrating how Indico excels in data extraction from invoice documents, financial records, rent rolls, and correspondences

### 4.1 Invoice Documents

**Challenge:** Invoices come in various layouts and formats, making manual data extraction time-consuming and error prone.

**Solution:** Indico's BERT-based models leverage contextual understanding to accurately extract information from invoices. Whether its vendor details, itemized lists, or transaction amounts, Indico captures the context and relationships within the document, ensuring precise data extraction. This capability streamlines the accounts payable process, expediting payment approvals and reducing the risk of discrepancies

### 4.2 Financial Documents

**Challenge:** Financial documents, such as statements and reports, often contain critical information dispersed across multiple sections

**Solution:** Indico's models excel in passing through financial jargon and complex structures. By understanding the contextual relationships within financial documents, Indico accurately extracts key data elements such as account balances, transaction details, and financial indicators. This ensures that banking professionals can swiftly analyze financial information, enhancing decision-making processes.

### 4.3 Rent Rolls

**Challenge:** Rent rolls, containing details about property leases and rental income, often require meticulous extraction of information for property management and financial reporting.

**Solution:** Indico's adaptability to document diversity allows for effective extraction of relevant data from rent rolls. The BERT-based models capture contextual nuances, extracting details like tenant names, lease terms, and rental amounts. This capability enhances the efficiency of property management processes within the banking sector, supporting accurate financial reporting and compliance.

#### 4.4 Correspondence documents

**Challenge:** Extracting specific information from diverse correspondences, including emails and letters, demands a nuanced approach.

**Solution:** Indico's contextual understanding extends to diverse forms of correspondence. Whether it's extracting customer details from emails or capturing relevant information from letters, Indico's BERT-based models discern the context and extract critical data elements. This capability streamlines communication processing, enabling banks to swiftly act on customer inquiries and requests.

### 5. CONCLUSION

In conclusion, this paper has extensively explored the transformative impact of Indico in automating document processes and data extraction within the banking sector. As we navigate the evolving landscape of financial services, the adoption of advanced automation tools becomes imperative, and Indico stands as a beacon of innovation with its powerful BERT-based models and versatile applications. In essence, Indico emerges not only as a solution to current challenges but also as a catalyst for future innovation in the banking sector. Its transformative impact, rooted in the precision of document handling, efficient data extraction, and adaptability to industry challenges, positions Indico as a cornerstone in the evolution of intelligent automation within the dynamic and complex realm of banking. As financial institutions continue their journey toward digitization and efficiency, the role of Indico remains pivotal in shaping the future landscape of banking automation.

### REFERENCES:

- [1]. Feiqi Huang, Miklos A. Vasarhelyi, Applying robotic process automation (RPA) in auditing: A framework, *International Journal of Accounting Information Systems*, Volume 35, 2019, 100433, ISSN 1467-0895
- [2]. Tiago Coito, Joaquim L. Viegas, Miguel S.E. Martins, Mariana M. Cunha, João Figueiredo, Susana M. Vieira, João M.C. Sousa, A Novel Framework for Intelligent Automation, *IFAC-PapersOnLine*, Volume 52, Issue 13, 2019, Pages 1825-1830, ISSN 2405-8963, <https://doi.org/10.1016/j.ifacol.2019.1>