BIZDATA

eZintegrations™
AI-Driven Document Understanding Solution
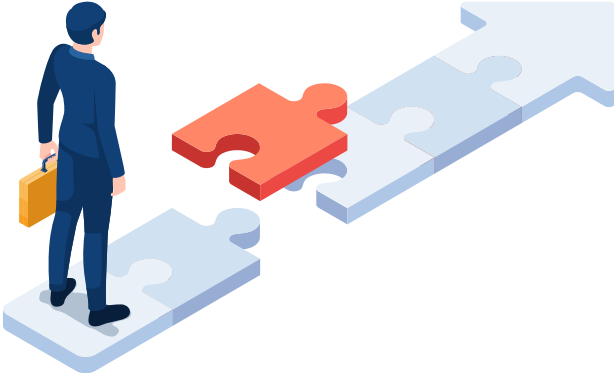
Release 2023

In today's data-driven world, organizations from all industries struggle to effectively extract insightful data from massive volumes of unstructured data, including paper-based documents and PDF files. From extracting critical information in clinical trials and financial reports to streamlining HR processes and enhancing compliance in the insurance industry, Document Understanding has the potential to reshape the way organizations operate.

This whitepaper seeks to clarify the relevance of document understanding integration and demonstrate how it may revolutionize a variety of sectors, including healthcare, banking, legal, and more. We speak about the expanding need to unleash the value of unstructured data and offer solutions for doing so. This paper will walk you through incorporating document understanding technologies into your workflow, whether you are a healthcare practitioner hoping to simplify patient record maintenance or a financial analyst looking to automate invoice processing.

Challenges

Content variations

The complex world of PDF files creates a significant barrier to data extraction. In the modern digital world, PDFs appear in a variety of forms, each of which hides a different set of complexity. Text-based PDFs, encrypted PDFs, and password-protected PDFs are all encountered; each requires a comprehensive strategy to reveal its secret contents. This digital maze demands navigating with a flexible toolset of techniques.

Formatting Issues

PDFs cover a wide range of visual content, from simple text documents to complex reports with tables, columns, a variety of typefaces, and changing font sizes. The difficulty of data extraction is increased by these complex changes. Maintaining the data's original format and meaning when dealing with PDFs that combine text, photos, and graphics becomes a difficult issue. It is crucial to maintain data integrity since any errors or misunderstandings might have an impact on applications that are used downstream.

Multiple Pages

The process of extracting data from PDF documents becomes noticeably more difficult when there are several pages present. PDFs frequently act as archives for large and thorough information, dispersing information over several pages. Extractors struggle to quickly navigate through these pages, identify important data pieces, and carefully put them together.

Real-time Data Extraction

Real-time data extraction is crucial in some applications, like financial or news reporting, where fresh PDFs are constantly being created. Real-time data extraction is a difficult technological challenge since it requires maintaining unchanging accuracy and efficiency.

Handwritten Content

The majority of information still comes in the form of handwritten notes and documents. It can be difficult to extract data from handwritten writing, especially when it lacks standardization. The different handwriting quality and styles might cause mistakes and inconsistencies throughout the extraction process.

Document Quality

Accurate data extraction can be hampered by poor document quality, which is frequently caused by scanning or photocopying. Smudged, faded, or distorted documents can present serious problems since the quality of the source document directly influences the precision of data extraction.
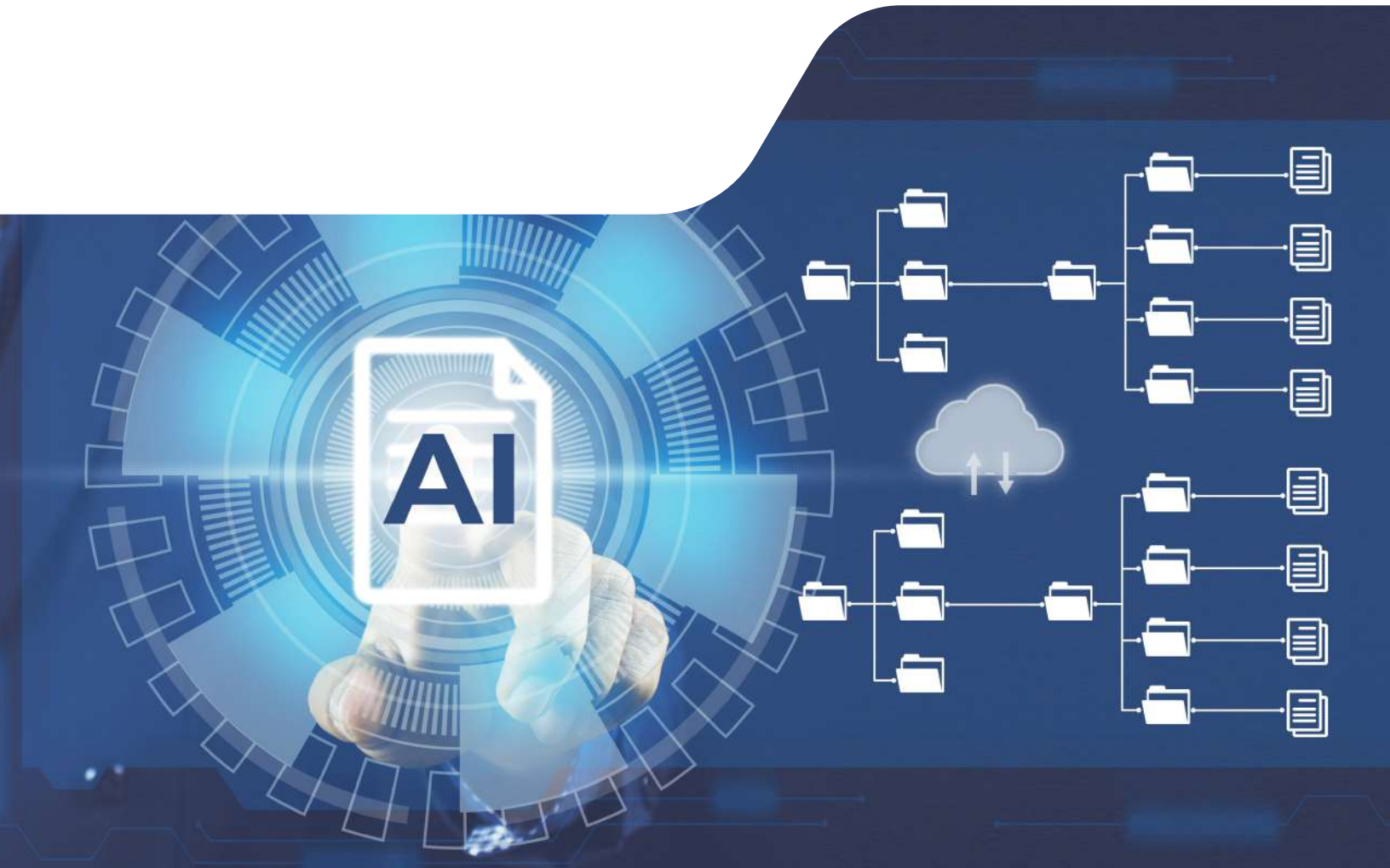
Sensitive Data Handling

Strict security procedures must be followed when handling papers containing sensitive or private information, such as bank records or data on individual identities. Understanding these papers is further complicated by the need to ensure compliance with data protection laws while extracting and processing them.

Scalability

The amount of paperwork that organizations must process grows as they expand. Scalability becomes problematic since conventional document understanding systems could find it difficult to effectively handle the increasing workload. It becomes extremely difficult to adjust to shifting needs without sacrificing performance.

Use Cases

Clinical Trials

The streamlined extraction of crucial information from a variety of clinical trial documents, such as patient records, informed consent forms, adverse event reports, and various essential documents, is a key strategy to achieve this goal in the field of healthcare and medical research. These studies, which serve as the foundation for medical progress, produce enormous and diverse datasets that call for effective administration. Automated data extraction provides a game-changing solution by utilizing cutting-edge technology like Natural Language Processing (NLP) and machine learning. It not only expedites the procedure but also improves data accuracy, enabling medical experts and researchers to decide on the best trial procedures and, eventually, quicken the creation of life-saving therapies.

Accounts Payable Automation

Effective financial management for firms depends on streamlining the invoice processing process. Automating the extraction of important information from invoices and smoothly integrating the gathered data into your enterprise resource planning (ERP) system is one extremely efficient method. Numerous benefits come from this process optimization, such as enhanced financial accountability, faster turnaround times, less manual labor, and higher accuracy.
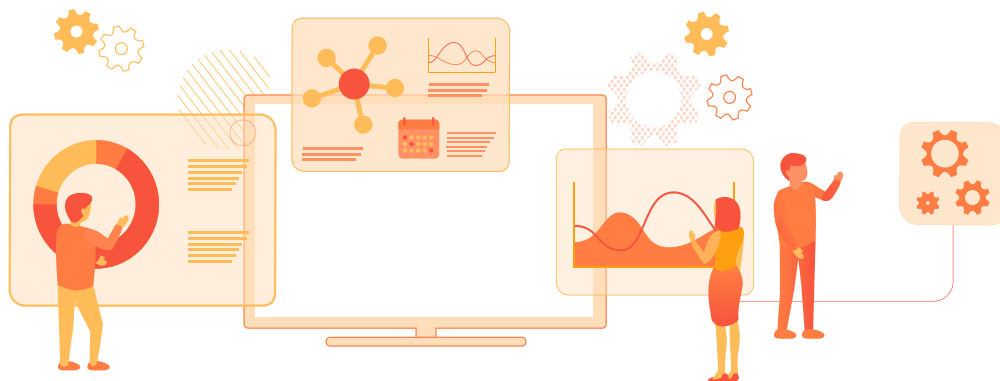
Automate Back-Office Tasks

By automating employee expense submission, laborious data input and paper effort are eliminated, enabling employees to report expenditures easily and get reimbursements more quickly. On the other side, automating new employee onboarding guarantees an organized and uniform procedure for bringing in new talent. Digital processes handle paperwork, training materials, and compliance documents, resulting in a speedier integration and higher employee satisfaction. The workplace may change as a result of these automation measures, which will boost overall productivity and free up labour for other important activities.

Digital Asset Management

Optimizing data indexing and retrieval procedures requires improving data management through the strategic integration of metadata. Data is enhanced with metadata, which acts as a contextual link and provides vital details about the source, organization, and usefulness of the data. Along with enhancing data quality and discoverability, enrichment also helps with compliance, governance, and data integration initiatives. Metadata facilitates effective cooperation, allows resource allocation optimization, and supports thorough data lifecycle management by supplying a consistent language for data description. Effective data management relies on the inclusion of metadata, which enables businesses to use their data for deft decision-making and superior operational performance.

Know-Your-Customer Verification

The automation of data extraction from multiple papers, including proofs of identification, proof of residence, and proof of income, makes it possible to streamline client onboarding. Businesses may increase productivity, decrease manual processing, and speed up the onboarding process by using automated systems that can extract and validate key information from these papers. This will result in a better experience for both customers and workers.

Supply Chain Management

Automating data extraction from crucial documents, including Purchase Orders, Sales Orders, and related paperwork, is necessary to efficiently streamline procurement procedures. Organizations can significantly improve accuracy, reduce manual labour, shorten procurement cycles, and improve inventory management by utilizing automated systems that can extract and process key information from these documents. This results in more effective and cost-effective procurement operations.

Insurance

For the insurance business to process claims faster and improve risk assessment, it is essential to streamline the extraction of critical information from insurance notes. Insurance companies may streamline operations, save processing times, increase accuracy, and ultimately deliver faster and more informed claims adjudication while also enabling more efficient risk assessment by implementing automated systems capable of quickly extracting crucial data from these notes.

Compliance and Regulatory Reporting

It is crucial to ensure compliance with industry laws, and one crucial tactic is to efficiently extract and organize data from compliance paperwork. Organizations can improve their ability to follow regulatory requirements, lower the risk of non-compliance, streamline reporting processes, and maintain a transparent and auditable record of their compliance efforts by implementing automated systems that can effectively extract and structure data from these documents. This will protect their operations and reputation.

Manufacturing

Automating the extraction of data from crucial documents including raw material specifications, batch processing papers, and packaging material specifications may dramatically improve manufacturing processes. By implementing automated systems that can effectively extract and process data from these documents, manufacturers are empowered to improve product quality, and production efficiency, ensure regulatory compliance, and enable data-driven decision-making, leading to more streamlined and competitive manufacturing processes.

Data Analytics

Utilizing the collected data for in-depth analytics is a crucial first step to allowing data-driven decision-making and acquiring insightful knowledge. Organizations may turn raw data into actionable insight by using sophisticated analytics tools and methodologies, and discovering trends, patterns, and opportunities. This encourages a competitive advantage, promotes strategic planning, improves operational efficiency, and enables data to become a valuable asset for the expansion and success of the organization.

Integration with Existing Systems

Maximizing the usefulness of retrieved data requires achieving seamless integration with your preferred data management and analytics solutions. Businesses may streamline their data workflows, speed up data processing, and provide real-time access to insights by guaranteeing compatibility and communication between their selected analytics platforms and data extraction tools. The total efficacy of data-driven plans and efforts will increase as a result of this integration, which equips users to make educated decisions, discover useful insight, and foster creativity.

Prescription Creation

The effective extraction of data from incoming prescription documents is necessary for automating the entry of prescription data. Healthcare providers can significantly improve accuracy, reduce manual data entry, minimize errors, improve patient care, and expedite prescription fulfilment by implementing automated systems designed to extract and process critical prescription details like patient information, medication details, dosages, and physician instructions. This results in a more effective and error-free prescription management process.

Process Manufacturing

For businesses that manufacture food, beverages, and other products using a process, effective data retrieval is essential for streamlining operations. Organizations that use a structured data management system may quickly access and use important data from a variety of sources, including packaging specifications, raw material specifications, compliance specifications, formulation specifications, and more. This all-encompassing strategy improves the control of product quality, regulatory compliance, and supply chain management, allowing for speedier decision-making and maintaining smooth operations across the whole production process.
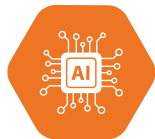
Key Features

Versatility

Our agile solution can easily navigate through a wide range of documents, including Purchase Orders, Sales Orders, Insurance Notes, Contract PDFs, Clinical Trial Documents, Recipes, Raw Material Specifications, Batch Processing Documents, Packaging Material Specifications, Compliance Documents, and hundreds of others. Its versatility has no boundaries, making it a vital tool for organizations in a variety of sectors. It successfully overcomes document variability to achieve unmatched efficiency and precision.

Data Source Integration

Our solution seamlessly integrates with your current infrastructure and works with a variety of data sources, including APIs, SFTP, SaaS platforms, NoSQL databases, conventional databases, and attachments coming from different applications. Organizations can simplify and automate the extraction of data because of this inherent flexibility, resulting in a new era of operational efficiency and accuracy.

Advanced AI Algorithms

Our technology excels in accurately analysing and comprehending document content thanks to cutting-edge AI algorithms, making it possible to extract crucial information quickly and accurately. The process of document comprehension is elevated to a degree of sophistication that is unmatched by its exceptional capacity to extract pictures while giving them contextual significance and deep knowledge.

Automated Data Entry

It skilfully extracts crucial information from receipts, identity papers, invoices, and a variety of other document types through seamless automation, classifying a wide range of documents. The operational efficiency is increased to previously unheard-of levels by this transformational capacity, which not only speeds up procedures but also greatly lowers the danger of human mistakes.

Content Repository System

By adding well-prepared metadata to your files, you may increase their worth. This enhancement forms the basis for better indexing and retrieval within data warehouses or digital asset management systems. Our technology not only improves the discoverability of valuable material but also streamlines document organization, allowing businesses to fully utilize their digital assets. Use metadata to your advantage to manage and utilize your data resources in a way that is more effective, intuitive, and knowledgeable.

Seamless ERP Integration

Effortlessly bridge the gap between document understanding and data integration by seamlessly storing the extracted data in the robust backend of leading enterprise resource planning (ERP) systems. Our solution seamlessly integrates with a spectrum of ERP systems, including Oracle E-Business Suite (EBS), Oracle Financials Cloud, Oracle NetSuite, Infor, SAP, Syspro, and an array of others. This integration not only ensures the accuracy and reliability of your data but also streamlines the flow of critical information throughout your organization's operational ecosystem, enhancing efficiency, compliance, and data-driven decision-making across the board. Say farewell to data silos and embrace the power of unified information management.

Data Integration

Create a seamless link between our solution and your current databases and systems, orchestrating a smooth flow of data. This interface ensures that the precise data gathered corresponds flawlessly with your current data architecture. By using the extracted data to improve decision-making, efficiency, and overall productivity, your organization will be able to function with agility and accuracy thanks to our dedication to expediting this crucial data interchange. Welcome to the era of unrestricted data synergy, where your current systems and our solution collaborate flawlessly to advance your company.

Image Extraction from PDF Files

Our method goes beyond simple text extraction by accessing pictures included inside PDF files to obtain them. This advanced function not only deepens data insights but also aids in an in-depth knowledge of the document's content. We provide organizations with a thorough understanding of both verbal and visual information by integrating picture retrieval smoothly, opening up new possibilities for data analysis and interpretation.

image_1	Id	name	Title	Product Description	Storage condition	Shelf Life	Part Number
	014B6I7G5QUREVSFXW56FJIBCGSZVJ3SUG	Product D.pdf	Product D light orange powered solid 2g in sachets	Product D light orange powered solid 2g in sachets	20 degree Celcius	48 months	*LOBK-161 R
	014B6I7G35ANF2PIRTJJDfWN5Q2MWVARQM	Product A.pdf	Product A tablets 300mg light blue almond shaped tablet	Product A tablets 300mg light blue almond shaped tablet with Brand and 4331 on one	below 30 °C	48 months	6195812345
	014B6I7GZWUT3DYYFURZF3lGRGF4RN6LTE	Product C.pdf	Product C tablets 200mg white round shaped tablet	Product C tablets 200 mg white round shaped tablet	below 35 degrees Celsius	24 months	*LOBK-141 R
	014B6I7G7H2IGJUEZCQ5HJVG4TR2B5Z86E	Product B.pdf	Product B Syrup 150 ml white colored thick consistency syrup	Product B Syrup 150 ml white colored thick consistency syrup	between 15 to 30 C	36 months	Product B Syr
	014B6I7G35S7XS7OSN2JCYLCGL7BXRBTAX	Product E.pdf	Product E capsule tablets 100mg red capsule tablets in a 25 count bottle	Product E tablets 100mg red capsule tablet with Brand and 4331 on one side	Store in 20 to 25 C	24 months	6195812386

Performance optimization

Our system is incredibly well-suited for businesses and organizations of all sizes since it was specifically designed to handle enormous amounts of data and documents. It gives you the ability to effectively manage, process, and extract insights from even the largest data warehouses, unleashing the unmatched potential for organizations of every size and level of complexity.
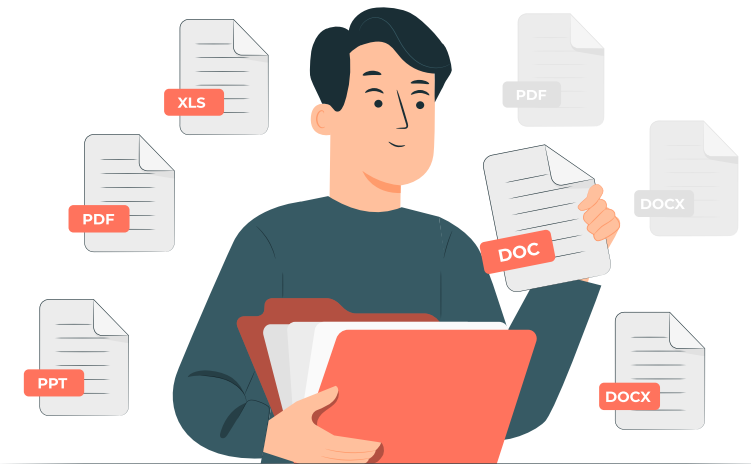
Data Security

We prioritize the security and confidentiality of your data. Our platform employs robust security measures to safeguard sensitive information.
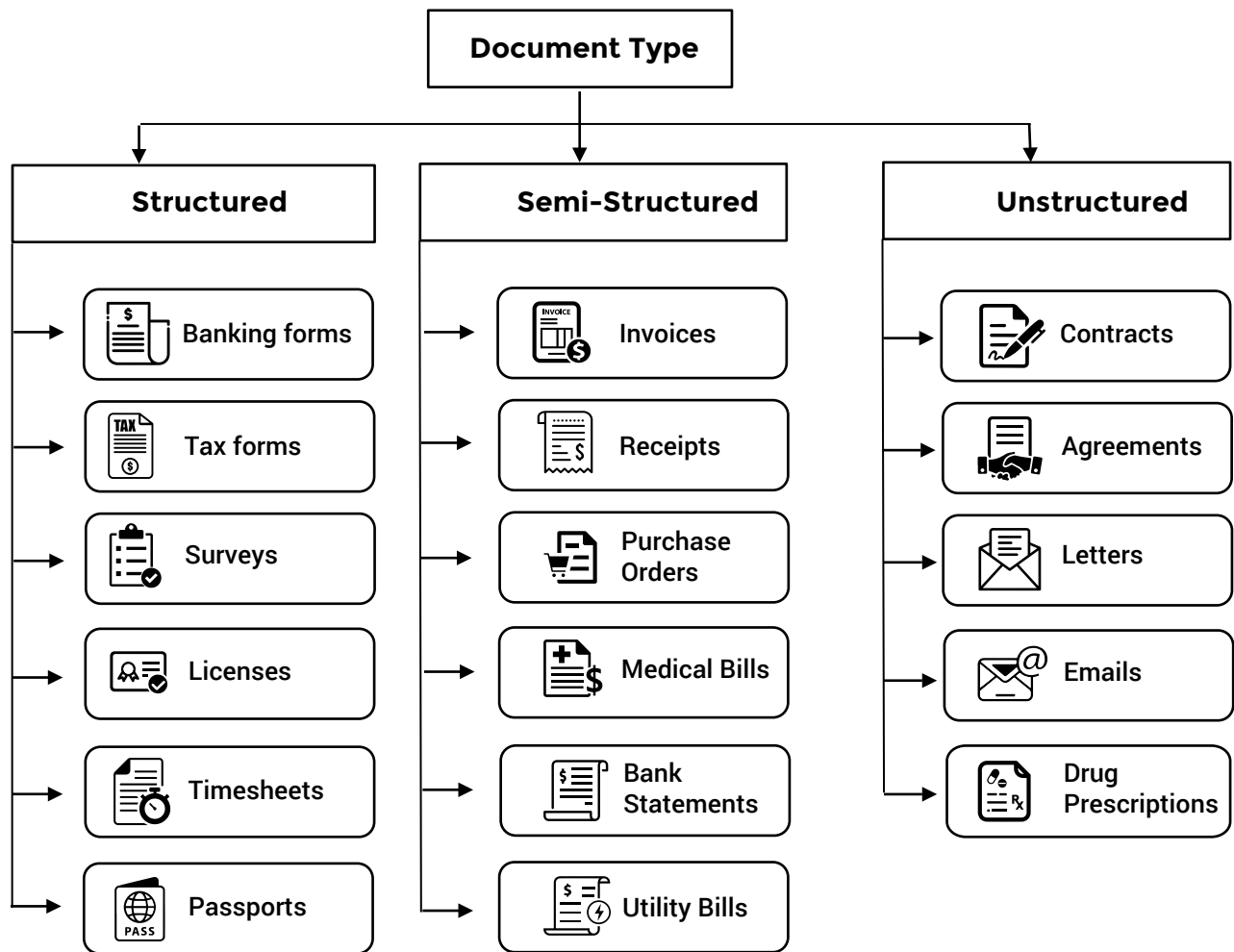

LABEL MASTER	
DOCUMENT NO.: LOBK-141	REVISION: 1.0
TITLE: Product C tablets 200mg white round shaped tablet	

The Part Number of the product is 6195812365

Item name and description- Product C tablets 200 mg white round shaped tablet with Brand on one side. These tablets are approximately 11.2 mm in length and 11.22 mm in width. Packaged with 10 tablets in a white strip with transparent film. Each strip has Aluminum film with moisture resistance, gas barrier, and light shielding properties.

Country of Origin printed on the label, reflects the country where the product was manufactured.

Britestock- The component has Product C tablets 200mg in white and round shaped with specification code LOBK- 141 and part number 6195812365 and quantity of item is 156.

Label-

- The component has Label, Strip with specification code LOMAT_LLB-1411 and part number 1411 and quantity of item is 156.
- The component has Carton with specification code LOMAT_LLB-1512 and part number 1512 and quantity of item is 156.
- The component has Insert with specification code LOMAT_LLB-1614 and part number 1614 and quantity of item is 156.

Tertiary Components-

- The component has Shrink Film, Polyolefin with specification code LO_TC_5555 and part number 5555
- The component has Shipping Case, Corrugated Box with specification code LO_TC_5558 and part number 5558 and quantity of item is 1.
- The component has Label, Bundle, 12 packs with specification code LO_TC_5589 and part number 5589 and quantity of item is 13.
- The component has Label Shipper, 156 pack with specification code LO_TC_5580 and part number 5580 and quantity of item is 2b.

Shelf life- The shelf life of the part is 24 months.

The part was manufactured on 01.2021 and expires on 01.2023

Labelling Instructions:

Label should be placed on top of bundle pack

Dimensions: Maximum size of label cannot exceed length and width of face in which the label is placed

Materials: Label Size should be suitable for its intended application

DOCUMENT STATUS: FOR APPROVAL   CONFIDENTIAL INFORMATION

LABEL MASTER	
DOCUMENT NO.: LOBK-171	REVISION: 1.0
TITLE: Product E capsule tablets 100mg red capsule tablets in a 25 count bottle.	

PART NUMBER:
6195812385

PRODUCT DESCRIPTION

Product E tablets 100mg red capsule tablet with Brand and 4331 on one side. These tablets are approximately 16 mm in length and 6 mm in width. Packaged with 25 tablets in a 60 mL white high density polyethylene bottles with one gram of desiccant and fiber packaging in each bottle. Each bottle is enclosed with a white continuous thread child-resistant polypropylene screw cap with induction sealed aluminum faced liner.

LABELLING DESCRIPTION

Country of Origin printed on the label, reflects the country where the product was manufactured

COMPONENTS	SPECIFICATION	PART NUMBER	QUANTITY
Britestock			
Product D 2g light powdered solid in sachets	LOBK-171	6195812385	90
Label			
Label, Bottle	LOMAT_LLB-3333	3333	90
Carton	LOMAT_LLB-3332	3332	90
Insert	LOMAT_LLB-3233	3233	90
Tertiary Components			
Shrink Film, Polyolefin	LO_TC_2222	2222	
Shipping Case, Corrugated Box	LO_TC_2228	2228	1
Label, Bundle, 12 packs	LO_TC_2229	2229	8
Label Shipper, 90 pack	LO_TC_2280	2280	2b

Document Understanding for different types of documents

```
                        Document Type
          ┌──────────────────┼──────────────────┐
          ▼                  ▼                  ▼
```

Structured	Semi-Structured	Unstructured
Banking forms	Invoices	Contracts
Tax forms	Receipts	Agreements
Surveys	Purchase Orders	Letters
Licenses	Medical Bills	Emails
Timesheets	Bank Statements	Drug Prescriptions
Passports	Utility Bills	

11

BIZDATA

Bizdata Inc.
www.bizdata360.com

Try eZintegrations™ for FREE and unlock the power of AI-driven Document Understanding.

Empower your organization with a robust, efficient, and secure solution to accelerate data extraction, improve accuracy, and enhance decision-making.

Contact us at
+1-650-283-1644
info@bizdata360.com