



Product details



ISSUED BY

Envirya Projects Private Limited

REPRESENTATIVE

Brig. Bishwajeet Ghosh, SM, VSM (Retd.)

Director, Catosoft Technologies Private Limited & Envirya Projects Private Limited

New Delhi - 110029

Bishwajeet.ghosh@envirya.in

+91 9560 847 186

AND

Kanishk Ghosh

Co-Founder, Envirya Projects Private Limited

New Delhi - 110029

kanishk.ghosh@envirya.in

+91 9953 572 420

CONTENTS

1. Introduction	3
2. Organization and Experience	3
3. The Value of En-Xtract.....	4
4. Benefits of En-Xtract.....	5
5. Use Cases for En-Xtract	6
A. Data extraction from admission forms.....	6
B. Digitization of old degree certificates / diploma certificates	6
C. Digitization of old mark-sheets	6
D. Digitization of old books and manuals.....	7
E. Digitization of invoices	7
F. Automated checking of OMR sheets	7
G. Translation of digitized data	7
6. Appendix	8
Table : Features and Subscription models of En-Xtract.....	8

1. Introduction

En-Xtract is an AI based product ([Please refer Appendix for detailed features](#)) that swiftly classifies and extracts critical data from documents of multiple layouts. It collects documents from input sources, makes them legible, and extracts data with precision and integrates this data with business's existing systems, while continuously improving extraction accuracy.

En-Xtract would be required for the implementation of digitization Certificates by Leveraging emerging technologies to maximize E-governance in companies and institutions.

2. Organization and Experience

Envirya Projects Private Limited is an Artificial Intelligence startup established in July 2019. The company is based in Gurgaon, India; registered with Startup India and is recognized as a startup by the Department for Promotion of Industry and Internal Trade with Certificate number DIPP40437.

Envirya is building a product named: "En-Xtract", which is an RPA tool for Intelligent Document Classification and Extraction. The company plans to provide its services to Government Bodies, MSMEs, Corporations, Startups, and everyday users. Opportunities to expand this service into Digitization, Insurance, TPAs, Banking and Finance, Invoice Generation, KYC services exist, and these avenues are actively being pursued.

Envirya's differentiation point comes with three simple principles:

- True collaboration with customers and partners
- Complete understanding of customers business
- Persistence in finishing the job whatever it takes.

Envirya Projects as an IT partner has a unique value in development, integration and implementation of Artificial Intelligence, Block-chain, Robotic Process Automation based projects in the following verticals: Insurance, Defense, Retail, Finance, Manufacturing, Health Care and Education. We have highly skilled team of Project managers, Consultants and Developers who are proven. We also have special expertise in custom software development – providing specialized products, IT services and custom end to end solutions to our enterprise customers.

Furthermore, we offer reasonable Service Level Agreements (SLA) covering most of the additional maintenance services. With our 24/7/365 days per year support, we are making troubleshooting as easy as possible.

Envirya Projects is a company which intends to be in front of the client's needs, deeply understanding clients' business and delivering and optimizing IT solutions from the initial phase of scoping the requirements, up to the final delivery, maintenance and continuous upgrade. Envirya always look at the things from the customer perspective, but also from customer's customer perspective to grasp their real business need for the IT solutions they deliver.

Our team operates with enthusiasm and flexibility. We are focused on our customers, being their reliable partner for IT services and solutions, working together on a partnership basis, helping them face the challenges of the modern times.

3. The Value of **En-Xtract**

- A. "En-Xtract" is a flexible solution to help automate the data entry process accurately and efficiently. En-Xtract reads your incoming unstructured documents and through the customizable platform, it intelligently classifies and extracts the information you need and routes your data and the original document to their destination.
- B. Our platform runs your source documents through an Optical Character Recognition (OCR) software and rules that can be written by Envirya or the companies and institutions from an easy to use, user-friendly front-end, without any knowledge of coding. This process not only reduces the time spent on manual entry, but also reduces human error typically caused by manual data entry and speeds up access to valuable discrete data so you can share, compare, report, and analyze the data. Our intelligent platform can recognize and categorize all types of incoming documents by content type and automate the extraction of relevant index data per document type.
- C. The manual indexing process is time consuming for your staff, and prone to errors. Changing your document handling and indexing workflow to a less manual process means your organization saves money, and your existing staff can focus only on quality control in the indexing process.

4. Benefits of En-Xtract

A. More automation, less manual work

The routing of documents can be fully automated from the moment it arrives to your organization until it is safely and accurately stored in the right place in your desired systems. The power of the En-Xtract platform allows workflows to be automated by intelligently indexing data.

B. Variety of documents - no templates or forms needed

What sets En-Xtract apart from our competition is that there is no need to know the structure or format of the documents or data in advance. By applying our proprietary technology, enhanced by machine learning, our software is able to recognize document types and specific data you want. There is no need to define in advance where specific data must be located within a document.

C. Automate all the workflows surrounding documents and unstructured data

The moment a fax, scanned document, or other unstructured data source enters your organization, it requires manual handling. En-Xtract can automatically identify document types and index data immediately to customize workflows based on document type.

D. Less manual handling means better security

When documents are printed and manually managed, your organization unintentionally runs the risk of exposing sensitive information. En-Xtract automatically sweeps up documents and data from secure network folders and then maintains your pre-defined security protocols to allow access to those documents for review before ultimately sending them securely to your desired storage location.

E. Speed of access to critical information

En-Xtract greatly reduces the time it takes for unstructured data or documents to reach the correct systems, with correct and consistent indexing data. For organizations with urgent document types, documents can be indexed and routed to desired area of the system within minutes. This reduction of manual processing empowers staff needing the information, to spend more time looking at documents instead of looking for documents. Allowing employees throughout the organization to make better, more informed decisions with the easily accessible data.

F. Indexing staff satisfaction

Start automating the more tedious and repetitive parts of your organization's document indexing process, allowing your staff to use their skills and knowledge for quality assurance checking or decision-making tasks. As volumes increase or staff turnover occurs, having an automated and consistent indexing system in place, alleviates pressure and allows quicker on-boarding of new staff.

G. Other staff satisfactions

Staff of all levels routinely express frustration with the increasing demands of data entry and searching for information. Expediting the flow of documents to your organization's document management system (DMS) or other relevant systems, En-Xtract will help eliminate these frustrations. Additionally, by automating the classifying, time spent on searching for documents is reduced and allows staff to easily access information they are looking for.

H. Unique document classification and advanced data extraction

Unique to En-Xtract, our automated document classification & indexing solution can route specific document types to our data extraction solution, for which a copyright is being filed. Extract can accurately identify and extract over 99.99% of the valuable data, which is currently being trapped inside documents. The system also enables to do Quality Check of the extracted data. If certain data is invalid, staff is alerted prior to it being filed as discrete information in the ERP or other backend systems. This results in a more complete database with both discrete data captured in the fields of the system and original document associated/linked to the data to reference if required.

I. Batch processing

En-Xtract allows for batch processing of documents where we can identify each document's type and extract data from them.

J. Other features of En-Xtract:

- Configurable in process making. Comes with common prefilled templates which can be modified based on the customer's requirements.
- By use of En-Xtract features, AI module can be integrated.
- Extraction of different languages : English, Hindi and German
- Translation to other languages of extracted data.
- Role Based Access Control (RBAC), for proper management of users for the companies and institutions. This will provide users in different hierarchies (within the organization/departments) different levels of control and access.

5. Use Cases for En-Xtract

A. Data extraction from admission forms

Required data can be extracted from the admission forms and can be fed to the system directly, thus abolishing manual data entry

B. Digitization of old degree certificates / diploma certificates

Old degree certificates / diploma certificates can be digitized

C. Digitization of old mark-sheets

An activity can be performed to digitize all the old mark sheets

D. Digitization of old books and manuals

Old books can be digitized. All the books available in the library can be digitize and thus an online library can be created wherein the students will not have to wait for the availability of the books if they are already issued to some other student. Any number of students can read the book simultaneously.

E. Digitization of invoices

An activity can be performed to extract the required data from Invoices (image or PDFs), for example: GSTIN, Invoice number, company name, address, line items, taxable amount, total amount, date etc., and thereafter integrating this with an ERP system of choice, such as an accounting software. These tasks can be performed in bulk for many invoices at a click of a button.

F. Automated checking of OMR sheets

Checking of the OMR sheets can be automated, then getting the results out in an instant.

G. Translation and Extraction of digitized data

Data extracted from documents whether in image or system generated form can be extracted and translated from the base language to a language of choice. Training of a language of choice would require approximately 15 days. **Currently data can be extracted in English, Hindi and German.**

6. Appendix

Table : Features and Subscription models of En-Xtract

	Standard	Pro	Enterprise	Description
	For companies that require basic data extraction with out-of-the-box accuracy. The processing happens on the cloud on a per transaction basis.	For companies that prioritize high accuracy and customization. The processing happens on the cloud with a license given to the company.	For enterprises that require an unlimited data extraction platform embeddable into their own software. If you require specific security and access measures for on-premise deployment.	
En-Xtract Features				
Validation UI	✓	✓	✓	En-Xtract's UI is optimized for rapid validation of extracted data. It can be customized according to your precise needs - additional fields, format checks, and more.
Custom Aliases and Regex	✓	✓	✓	Customise the aliases and regular expressions according to your needs to pick up any new formats in your documents
MagicGrid for tables	✓	✓	✓	MagicGrid is a smart tool for visually capturing table-like data - this happens automatically when paired with a Dedicated AI engine
Mass document scanning	✓	✓	✓	Upload multiple documents in a single scanned file and En-Xtract will detect page breaks
API Access		✓	✓	Advanced API access including importing, exporting, user management, schema customization, and other metadata.
Quality Assurance		✓	✓	Allows you to check all documents and make edits if necessary
Classification of Document Category			✓	Ability to classify different categories of a document into their respective types
Extension environment			✓	Customize En-Xtract by implementing simple microservices that can automatically process extracted data according to your business rules.
Dedicated AI (1 schema, recurring training)			✓	AI Engine trained on customer-provided data for their specific demands
Organization creation privileges			✓	Allows the creation and manipulation of organizations under your superuser account.
High-scale requirements				
# of processes	3	3	Infinite	If you require specific security and access measures our sales is ready to help. Processes determine the workflow and lifecycles of uploaded documents and the data that is extracted from within them.
Multiple input source	✓	✓	✓	Inputs from various data sources can be taken like emails, ZIP, SFTP folders etc.
Pre defined templates for standard invoices	✓	✓	✓	For most common invoice formats, pre-created templates will be provided - these will be customisable
Output format	✓	✓	✓	Output can be provided in the desired format - CSV, JSON, Word, Notepad etc.
Unlimited users		✓	✓	
Template customization		✓	✓	Customization of templates can be done as per needs
Pluggable extraction algo			✓	Multiple AI algos can be provided to test which algo is giving most accurate result for the document from which data needs to be extracted
High-priority document processing			✓	Process priority management of the same process
Analytics - dashboards			✓	Various dashboards can be provided like- number of documents remaining in the package, userwise consumption
API calls higher frequency usage			✓	Depending on API calls per second
Support				
Self-serve knowledge base	✓	✓	✓	
NDD email support	✓	✓	✓	
Access to solution engineer			✓	
24/7 phone support			✓	
Security				
VPN connection to customer systems			✓	
Single-tenant cloud infrastructure			✓	

ENVIRYA EN-XTRACT



**FOR INTELLIGENT
DATA EXTRACTION**

Envirya [ΣN]-Xtract

Most of the critical business data remains hidden in ever-increasing volumes of paper or electronic documents. You, as a business leader, need a way through which your users can mine valuable information held within a document and make well-informed decisions.

[ΣN]-Xtract, swiftly extracts critical data from documents of multiple layouts. It collects documents from input sources, makes them legible, and extracts data with precision and integrates this data with your existing systems, while continuously improving extraction accuracy.

DOCUMENTS & IMAGES

1. INGEST

ECM REPOSITORY



MULTIPLE DOC FORMAT



MULTI-FUNCTIONAL
PRINTER & SCANNERS

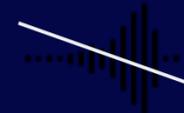


2. CLASSIFY & PROCESS

ORIENTATION &
SKEW CORRECTION



BACKGROUND
NOISE REMOVAL



CLASSIFY
AS PER
USER DEFINED BUCKETS



3. EXTRACT & QA

USER FRIENDLY
INTERFACE



MULTIPLE LAYOUTS



ONE CLICK EXTRACTION



4. LEARN

DATA EXTRACTION
ANALYTICS

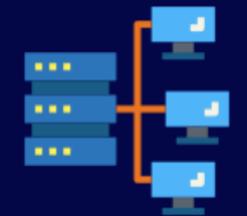


DOCUMENTS LAYOUT
LEARNING



5. INTEGRATE

UI BASED
INTEGRATION



WHY

-XTRACT

The Need:

Faster and better decision making and cost-reducing workflow efficiencies

The Problem:

Manual data entry of high-volume documents with complex or inconsistent layouts from external sources

The Solution:

Quick, accurate, and cost-saving information extraction

The Framework:

Diverse content recognition through machine learning, rules-based data models, and natural language processing





INGEST

DOCUMENT INGESTION -

Ingest documents and images in multiple formats, from different content sources

Multiple formats supported, such as PDF, Png, Jpeg, TIFF, MS-word, MS-excel, RTF, and Email

Different content sources include multi-functional printers, scanners, enterprise applications, and repositories

PROCESS & CLASSIFY

DOCUMENT PRE-PROCESSING

Improve quality of single page or multiple page scanned documents in real time with automatic detection, correction of image distortions & classification

COLOR AND GRAYSCALE IMAGING

AUTO-CORRECTION
OF DOCUMENT ORIENTATION

SKEW CORRECTION

AUTOMATIC BACKGROUND NOISE
REMOVAL

CLASSIFY DOCUMENTS INTO
BUCKETS





Extract & QA/QC

DATA DEFINITION

Leverage user friendly interface for defining multiple document types with varying layouts

MULTIPLE DATA FORMATS EXTRACTION

Extract complex data formats from documents of different layouts into a searchable database

Table data recognition from multiple table structures

Hand-written text extraction

Predefined data formats with regular expression validation, such as Address, Date, Invoice Number etc.

WEB-BASED VERIFICATION

Assist workers with automated interfaces for speedy, extraction, verification of unrecognized and recognized data

One-click extraction with cursor pointing for undefined labels

High-speed Quality assurance and Quality check through a user-friendly User Interface

Learn

Document Layout Learning

Continuously enhance accuracy in data extraction by learning from processed documents

Generate labels, data, types, and locations by learning from manual data extraction of unrecognized document

Data Extraction Analytics

Monitor extraction process and accuracy rates. Get visual indications to find data inaccuracy and make corrections

Verify extracted data via system generated trend reports

Check accuracy of extracted data via system generated confidence rate

Integration

INTEGRATION FRAMEWORK

Extract data from large volume of documents across your business processes



Scalable extraction at production stage

UI based integration of En-Xtract

Run multiple extraction services simultaneously

Push extracted data into your integrated ERP system in the required format



FOR INDUSTRY SPECIFIC DOCUMENT TYPES

Cross Industry

Invoices
Purchasing Order
Appraisal Forms
Sales Order
Expense Forms
Surveys
Questionnaires
Salary Slips
Answer Sheets
On-Boarding Forms
Claims

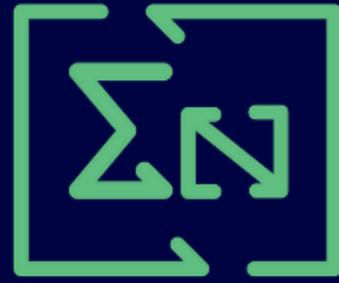
Banking & Insurance

Loan processing form
Cheques
Bank statements
KYC form
Account opening form
Credit Card applications
Insurance Forms
Proposal Forms
Medical Documents
Lab Reports

Government

Application forms
Passports
Payment Receipts





ENVIRYA

Envirya Projects Pvt. Ltd. is a vendor/provider of Business Process Management (BPM), Enterprise Content Management (ECM), Document Management System (DMS), Workflow and Process Automation Software.

contactus@envirya.in

www.envirya.in

Sales Inquiry : +91-9953572420