Intelligent Document Classifier

Release Notes

Version: 2.1

Release Date: 16th August 2023

Disclaimer

This document contains information proprietary to Newgen Software Technologies Limited. The user may not disclose or use any proprietary information or use any part of this document without written permission from Newgen Software Technologies Limited.

Newgen Software Technologies Limited makes no representations or warranties regarding any software or the contents or use of this manual. It also specifically disclaims any express or implied warranties of merchantability, title, or fitness for any particular purpose. Even though Newgen Software Technologies Limited has tested the hardware and software and reviewed the documentation, it does not guarantee or imply that this document is error-free or accurate regarding any particular specification. As a result, this product is sold as it is, and the user, the purchaser, is assuming the entire risk as to its quality and performance. Further, Newgen Software Technologies Limited reserves the right to revise this publication and make changes in its content without any obligation to notify any person of such revisions or changes. Newgen Software Technologies Limited authorizes no Newgen agent, dealer, or employee to make any modification, extension, or addition to the above statements.

Newgen Software Technologies Limited has attempted to supply trademark information about company names, products, and services mentioned in this document. The trademarks indicated below were derived from various sources.

Copyright © 2023 **Newgen Software Technologies Ltd.** All Rights Reserved.
No part of this publication may be reproduced and distributed without the prior permission of Newgen Software Technologies Limited.

Newgen Software, Registered Office, New Delhi
E-44/13
Okhla Phase- II
New Delhi - 110020
INDIA
Phone: +91 1146 533 200
info@newgensoft.com

Table of contents

1 About this document 4
1.1 Revision history 4
1.2 Intended audience 4
1.3 Documentation feedback 4
2 Release summary 5
3 Key highlights 5
4 Supported platforms 6
5 System requirements 6
5.1 Hardware requirements 6
5.2 Software requirements 7
6 Upgrade impact 7
7 Deployment steps 7
8 Rollback plan 7
9 Integration strategy and their dependencies 7
10 Software deliverables 8
11 Docker deliverables 8
12 Known constraints 8
13 Release artifacts 9
14 Version number 9
15 Build number 9
16 Release dates 9
17 Product development location 9

1 About this document

These release notes provide information on the features, functionalities, enhancements, and known limitations of the Intelligent Document Classifier (IDC) 2.1. These release notes also include system requirements for installing IDC 2.1.

1.1 Revision history

Revision date	Description
August 2023	Initial publication

1.2 Intended audience

These release notes are intended for new and existing customers of the IDC software, Newgen's partners, implementation, sales, and the pre-sales team.

1.3 Documentation feedback

To provide feedback or any improvement suggestions on technical documentation, write an email to docs.feedback@newgen.co.in.

To help capture your feedback effectively, share the following information in your email:

- Document name
- Version
- Chapter, topic, or section
- Feedback or suggestions

2 Release summary

IDC 2.1 is a subsequent release of Newgen's document classification solution, building upon the capabilities of its predecessor - IDC 2.0. It uses a classification approach, capable of handling various document types, achieved by automatically recognizing diverse structural and textual features within the document. Classification can be performed based on structural features (Document Image Classification), textual features (Document Content Classification), the document's layout (Object-Based Classification), or using any combination of these models.

It is implemented as a Python-based web API, where a document is provided as input, and deep learning algorithms are utilized to classify its pages. This release encompasses all the features of the previously launched IDC versions, thus, enhancing the overall document classification experience.

3 Key highlights

The following are the key highlights of the IDC 2.1 release:

- Enhanced security compliances with the addition of security patches.
- Upgraded libraries and Python version for improved performance.
- Introduction of a containerized version for convenient deployment.

4 Supported platforms

This section lists the supported server and client environments.

Server environment

IDC 2.1 supports the following server environments:

Operating system	Windows 2016/2019 (64-bit) Enterprise Server
Database	MongoDB 4.x
Application server	CherryPy 18.6
Language	English (United States)

Client environment

IDC 2.1 supports the following client environments:

Supported browser	• Microsoft Edge 113 and later • Google Chrome 113 and later • Mozilla Firefox 113 and later

5 System requirements

This section lists the minimum hardware and software requirements for installing IDC 2.1.

5.1 Hardware requirements

The table below lists the minimum hardware requirements for the server machines:

Server	• 64-bit machine • 4 core dedicated processor (with at least 8 virtual cores) • 16 GB RAM and above • 50GB hard drive • AVX instructions set support

NOTE:
Refer to the server sizing sheet or performance report to determine the precise hardware requirements necessary to manage the peak volume of processing within the expected time limit.

5.2 Software requirements

The table below lists the minimum software requirements for the server and client machines:

Server	• Refer to the Supported platforms section to set up a combination of the operating system, and database server. • CherryPy 18.6 Web Server
Client	Refer to the Supported platforms section.

6 Upgrade impact

IDC 2.1 does not support the old models that were compatible with IDC 2.0 or earlier versions. To ensure compatibility with any Newgen or external application, ensure that you perform the integration with the latest binaries of IDC 2.1.

7 Deployment steps

To deploy IDC 2.1, refer to the *Intelligent Document Classifier 2.1 Installation and Configuration Guide*. For containerization-based deployments, refer to the *Intelligent Document Classifier 2.1 Deployment Guide*.

8 Rollback plan

Upon the execution of the rollback script, the IDC system along with its docker images get removed from the host server machine. However, the user datasets in the *Mount* folders remain unaffected.

9 Integration strategy and their dependencies

To ensure compatibility with any Newgen or external application, ensure that you perform the integration with the latest binaries of IDC 2.1.

10 Software deliverables

The following software deliverables are available with IDC 2.1 release:

- **Manuals** (contains IDC-related documents)
  - Installation and Configuration Guide
  - Deployment Guide
  - Developer Guide
  - Best Practices Guide
- **Release Notes**
- **Setup** (binaries to install IDC 2.1)

11 Docker deliverables

The following docker deliverables are available with IDC 2.1 release:

- Docker images for IDC inference API and MongoDB
- Configuration files
- Folder structure for Content, Image and Object models (*Mount* folder)

12 Known constraints

The following is the list of the known constraints in the IDC 2.1 release:

- The accuracy of the analytics model depends on the image quality and variations for the image-based classification or object-based classification.
- The input documents must adhere to Newgen's specified capture and quality standards.
- Being a machine learning-based solution, it is important to note that achieving 100% accuracy is not guaranteed. However, it is possible to observe good classification even on visually poor-quality images, while encountering failures on visually good-quality images.
- If the skew of an image exceeds +/- 5 degrees, it may adversely affect the classification accuracy.
- The performance of the product might not be optimal when dealing with document variations that were not used during the model training phase. For example, a model developed for a specific state's driving license may not be suitable for other state driving licenses.
- The execution timings per page have extended by 0.2 seconds on an average due to the typical outcome of implementing security patches and enhanced compliance measures.

13 Release artifacts

The release artifacts are as follows:

S. N.	Enclosed	Yes / No / NA
1.	Installation guide (describing installation procedure for the network as well as standalone version; dependency of software on the operating system and other software; structure and paths of various *.*ini* files, minimum software, and hardware requirements)	Yes
2.	New function listing	No
3.	Modified functions	Yes
4.	Interface modifications	NA
5.	Problem status (for all the pending bugs, give bug number and comments, and provide further details).	No
6.	COTS components (If Any)	NA
7.	Others (if any, for example, Freeware, open-source, executables).	NA

14 Version number

Intelligent Document Classifier 2.1

15 Build number

3965

16 Release dates

Following are the release dates:

- **Product release date** – 16th August 2023
- **Market launch date** – 30th August 2023

17 Product development location

Newgen Software Technologies Limited,
9th Floor, Building No.7, Candor TechSpace,
IT/ITES SEZ of M/s Seaview Developers Pvt. Ltd.,
Plot No.20 & 21, Sector-135,
Noida, Uttar Pradesh – 201304
E-mail: helpdesk@newgensoft.com