Information Assurance Disclosure of Xerox® Proofreader

1. Introduction
Xerox® Proofreader for performing a full grammar, spelling, style, punctuation and plagiarism check is a workflow solution that connects Xerox® Multifunction Printers (MFP) to the WhiteSmoke proofreading solution. Printing, Proofing, and viewing corrections is easy and convenient from Xerox® MFP devices without the need of servers, 3rd party scan equipment, or manual processing of test and results. This reduces time and cost with ensuring privacy and security.

1.1. Purpose
The purpose of the Information Assurance Disclosure (IAD) is to disclose information for the proofreader solution with respect to device security. Device security, in this context, is defined as how data is stored and transmitted, how the product behaves in a networked environment, and how the product may be accessed, both locally and remotely. This document describes design, functions, and features of the Xerox®Proofreader relative to Information Assurance (IA) and the protection of customer sensitive information. Please note that the customer is responsible for the security of their network and the Xerox®Proofreader does not establish security for any network environment.

The purpose of this document is to inform Xerox® customers of the design, functions, and features of the Proofreader relative to Information Assurance (IA).

This document does not provide tutorial level information about security, connectivity or Xerox®Proofreader features and functions. This information is readily available elsewhere. We assume that the reader has a working knowledge of these types of topics.

1.2. Target Audience
The target audience for this document is Xerox® educational institutions and customers concerned with IT security.

It is assumed that the reader is familiar with the Proofreader app; as such, some user actions are not described in detail.

1.3. Disclaimer
The content of this document is provided for information purposes only. Performance of the products referenced herein is exclusively subject to the applicable Xerox® Corporation terms and conditions of sale and/or lease. Nothing stated in this document constitutes the establishment of any additional agreement or binding obligations between Xerox® Corporation and any third party.
2. Description and Details

2.1. Overview
The Xerox® Proofreader app provides three primary workflows:

- Grammar Report
- Plagiarism Results
- Review & Print/Email results

Each workflow facilitates a combination of the steps:

- App hosting
- Authentication
- Selection
- Scanning
- Printing
- Emailing

2.2 App Hosting
The Xerox® Proofreader app consists of two key components, the device app and the API. The device app is a ConnectKey / EIP webapp and the API is a REST API, both are served by webservers in the cloud.

2.3. Authentication
Authentication consists of 2 key steps.

**User creation:** A new user is required to link/connect their Xerox® Proofreader account to the device. Adding themselves requests their username and password. These values are passed through the Xerox® Proofreader Application Programming Interface and forwarded onto the Cloud API. An OAuth login token is returned to the device which is used for further interactions. At this time, the user is prompted for a password to facilitate convenient login.

The user’s token, along with their name as it appears in the database and email address (to ensure no duplicate entries) is stored on the Xerox® Proofreader cloud server.

**User login:** when the user has linked their account to the Proofreader app. Once validated, the Proofreading app will refresh the stored token, via the Proofreading-API, allowing the user the ability to use the app for the given session.

The token is updated on the local device once refreshed.

2.4. Selection
At various steps in the application the user may be prompted to make selections, these include Plagiarism results, Grammar, Style, and Spelling. These lists are dynamic and driven by API calls to the Proofreading-API with the user’s OAuth token.

2.5. Scanning
When scanning documents for Proofreading or Plagiarism, documents are scanned and submitted to the Proofreading-cloud. Due to the nature of the Xerox® EIP scanning workflow design, the users OAuth
token is temporary persisted (encrypted) in the Proofreading-API database. As the scan is received by the Proofreading-API it is forwarded onto the OCR-API along with the temporarily persisted OAuth token. Once the scan process is complete this token is deleted from the database. If the scan process is interrupted, the token value will be removed by the Proofreading-API privacy workflow which removes old records. This workflow is executed multiple times per hour.

2.6. Printing
When printing tests or reports, the request is sent to the Proofreading-API along with the user’s OAuth token. Data is then forwarded to the grammar and plagiarism API and the relevant reports are generated. These reports are sent back to the Proofreader-API database where they are temporarily persisted. URL links for those reports is sent to the Xerox® device for use with Pull-Print. These URLs are short live and removed upon expiry.

2.7. Emailing
In conjunction with printing, the user can choose to email reports to themselves. In both cases the request to email is sent to the Proofreader-API along with the OAuth token. Data is then forwarded to the grammar and plagiarism API and the relevant reports are generated and the API sent the reports as download links to the customer email.

2.8. SNMP & Device Webservice Calls
During standard usage of the Proofreader app, local calls to SNMP are initiated to pull relevant details such as device language and paper size preferences. The initiation of scan, print, and the usage of internal graphical components are handled through these local webservice calls.

3. Security
3.1. Hosting
The Xerox®Proofreader-API and the Portal app are hosted on the Amazon Web Services Network. Amazons data center operations feature comprehensive information security policies and processes using standardized industry control frameworks, including ISO 27001, SOC 1, and SOC 2.

For a full description, please follow the link: https://aws.amazon.com/security/

3.2. Secure Web Communications
All web communications between servers and Xerox® devices are encrypted using HTTP Secure (https).

3.3. Encryption
When required, the user’s OAuth token is persisted on the device and in the cloud (temporarily). While at rest, this token is salted and encrypted using AES-256. Token encryption and decryption is done on the server. The Xerox® device has no internal means to decrypt this value.
3.4. Data
The only data that is being stored on WhiteSmoke servers is the actual document data itself in various formats (PDF, DocX). The data is being stored on our database which is located in the USA and is run through amazon. Please note the data is being stored for a maximum of 48 hours before being completely erased from our database and all documents are being encrypted with the latest encryption methods to provide maximum security. (The encryption method that is used on all documents is AES 256 which is applied to any documents that are received to our database)

**Flow of data**

<table>
<thead>
<tr>
<th>Printer App - when someone scans a document the printer app converts it to base 64 encoding and sends the document to the WhiteSmoke API</th>
<th>Portal - when someone type text or upload a document using the portal the data is encoded with base64 encoding and send the data to the WhiteSmoke API</th>
</tr>
</thead>
<tbody>
<tr>
<td>The API gets the document data and encrypts it using AES encryption (using different encryption key for each user) and saves the data to the WS database</td>
<td></td>
</tr>
<tr>
<td>After data has been saved it is forwarders to an external service called ABBY for OCR conversion (from PDF to DOCX).</td>
<td></td>
</tr>
<tr>
<td>When converted data has been received from ABBY we encrypt and store it again on the database.</td>
<td></td>
</tr>
<tr>
<td>The API then sends the converted document to the WhiteSmoke engine for grammar/plagiarism checking.</td>
<td></td>
</tr>
<tr>
<td>Result of the WhiteSmoke check is encrypted and stored on the database (up to 48 hours). After 48 hours data is deleted from the database</td>
<td></td>
</tr>
</tbody>
</table>

4. Ports
4.1. App & API
The Xerox®Proofreader app and Portal require that the device is able to communicate over port 443 outside the client’s network.