iDNA Applications Data Collection and Access

This page describes what iDNA reads from your HCL Notes and Domino environment, which access it needs for that, and which connections it uses. It applies to every license, with the differences on a trial license outlined in one section at the end. It describes iDNA Applications 3.5.0 and later.

For what happens to the data after iDNA has read it, see iDNA Applications Data Flow in Detail. It covers storage, processing, who sees what, and what leaves your network.

- Overview
- The Collection ID and the Access It Needs
- Connections
- What iDNA Reads, by Source
- Is Document or Mail Content Read?
- Personal Data
- On a Trial License
- Related Topics

Overview

iDNA Applications runs on a Linux host in your network (the appliance). It reads from the Domino servers you select, with one Notes ID that you provide: the collection ID.

Summary

- **Read-only towards Domino.** iDNA creates, changes or deletes no documents, designs or ACLs in your databases and runs no agents. The only console command it sends is `show config *`, which lists the server's notes.ini settings.
- **Connections start at the appliance:** Notes RPC (TCP 1352) to your Domino servers, and HTTP or HTTPS to a Domino web server if web logs are read as text files.
- **From system databases:** the Domino Directory, server documents and notes.ini settings, `log.nsf`, `catalog.nsf` and `domlog.nsf`. With the optional packages, also the DLAU and MarvelClient Analyze databases.
- **From application databases:** the design and source code of Focus Databases and templates, the database's record of user activity, Profile Documents in full, and document metadata. No other document content, and no attachment contents.
- **Personal data:** Notes user names, directory data about people and groups, IP addresses and client details from web logs and MarvelClient, mail addresses from the router's log. Stored on the appliance and pseudonymized in the analyses by default.
- **Data stays on-prem:** collected data is processed and stored on the appliance in your network. It only leaves when you choose: results for your AI assistant if you enable the MCP server, alerts and report subscriptions by email, and support bundles or license requests an administrator sends.

The Collection ID and the Access It Needs

iDNA reads everything with the collection ID, which needs the access below. The reference list is in System Requirements, Access and Permissions.

Access	Where	Used for
Access to the server	Every HCL Domino server in scope	Any collection from that server
Reader	The Domino Directory (`names.nsf`) on one server per domain, and by default the directories configured in Directory Assistance	The Domino Directory package
Reader	`log.nsf` and `catalog.nsf` of every server; `domlog.nsf` where web logging to it is enabled	The Server Basics, Database Catalog, Notes Client Activity and Web User Activity packages; the optional Domino Event Log and Mail Routing Activity packages
Full Remote Console Administrator	Server document of every server in scope	The console command `show config *` and the console part of the server access check
Designer, or Full Access Administration	Designer in the ACL of every Focus Database and template; or Full Access Administration in the server document	Design analysis. Where the ID has Full Access Administration, iDNA uses it by default. Reader access is the minimum for a database to be opened at all.
Reader	The DLAU result database; the MarvelClient Analyze database	The optional HCL DLAU and MarvelClient Analyze packages
Reader, or Full Access Administration where the ID has it	Each database a content scan is started for	Content analysis

- **Full Access Administration** lets the ID open every database on that server whatever its ACL says, mail files included. iDNA uses it to read designs. Designer access per database is the narrower alternative.
- **Servers of other organizations:** iDNA collects only from servers certified by an organization named in your license. Other servers it finds are listed as unlicensed and not collected. Multiple certifier names can be included in a license.
- **Server prerequisites** (the Catalog and Statlog tasks, session logging, the web log format) are listed in System Requirements.
- **The ID file and its password** are stored on the appliance, so that iDNA can collect without anyone logged in. See Data Flow in Detail, Storage on the Appliance.

Connections

From To	Protocol, port	Purpose	When
Appliance Domino servers	Notes RPC, TCP 1352	All collection from Domino	Always
Appliance Domino web servers	HTTP or HTTPS, the port you configure	Reading web logs as text files	Only if the Web User Activity package is set up to read text web logs over HTTP
Appliance your DNS servers	DNS	Resolving the Domino servers' host names	Always

Your Domino servers open no connection to the appliance. For browser access to iDNA, LDAP logins, email, AI clients and every connection that leaves your network, see Data Flow in Detail, Who Sees What and What Leaves Your Network.

What iDNA Reads, by Source

iDNA collects in three ways:

- **Content packages**, set up per server in the admin interface (see Admin Interface and Config Manager). Each package reads one kind of data from the servers it is active on.
- **Design analysis** of Focus Databases and templates, with the activity record, Profile Documents and content age of each Focus Database.
- **Content analysis**, started by an iDNA user for a single database.

Content Packages

Package	Read from	What is read	Personal data
Domino Directory	`names.nsf` on one server per domain; directories in Directory Assistance	Person documents, group documents (with members and owners), mail-in database documents	Names, organizational data and mail and internet addresses of people; group memberships
Server Basics	Every server in scope	Server documents, connection documents, notes.ini settings (through `show config *`), server statistics (hourly)	Administrator names in server documents
Database Catalog	Every server in scope; `catalog.nsf`, `log.nsf`	The list of databases with their properties (title, file path, replica ID, template, size, dates), catalog entries, the number of ACL entries per access level, database activity counts	None: ACL entries are counted, not stored by name
Notes Client Activity	`log.nsf`	Session records: who opened which database on which server, when, with read and write counts (every 2 hours)	Notes user names
Web User Activity	`domlog.nsf`, or text web logs over HTTP or HTTPS	Web requests: user name, client address and host name, user agent, requested URL, time. Cookies and referrers are not stored.	User names, IP addresses, host names
MarvelClient Analyze (optional)	The MarvelClient Analyze database	Client records: Notes user name, home server, Notes and MarvelClient versions, operating system, operating system user name, IP address	All of these
HCL DLAU (optional)	The DLAU result database	Its documents in full, rich text included	User names and the licensing data DLAU recorded about them
Domino Event Log (optional, licensed separately)	`log.nsf`	Events and security events: time, type, severity, server task and message text	Names that appear in event messages
Mail Routing Activity (optional, licensed separately)	`log.nsf`	The router's transfer and delivery events: sender, recipients, size, time	Mail addresses

Design Analysis

Design analysis covers the **Focus Databases** and one replica of each template. Focus Databases are the databases of type Application and Mail-in, as classified from the catalog. You can adjust the classification with recategorization rules. Design analysis runs one database at a time. It re-reads a design when it has changed, by default weekly on Saturday.

Source	Read from	What is read	Personal data
Design	Focus Databases and templates	The full design, exported as DXL. It contains every design element and its code (LotusScript, Formula language, JavaScript, Java, XPages) with names, comments, signers, last modifiers and dates. It also contains the database properties and the ACL with its entries, roles and log. iDNA keeps the export on the appliance as a compressed file. No data documents.	Names of signers, modifiers and ACL entries; anything developers wrote into the code, such as names, addresses or credentials
User activity record	Focus Databases and templates	The activity that Notes records in each database: user names, dates, reads and writes	Notes user names
Profile Documents	Focus Databases	Every Profile Document in full: all items and their values, its modifiers and revision dates	Whatever the application stores in its Profile Documents; modifier names
Content age	Focus Databases (content-age module)	The number of documents and their creation and modification dates, stored as counts per month. No item values, no form names.	None

More on content age: Content Age and Data Retention Analysis in iDNA Applications.

Content Analysis

Content analysis has two parts. Content age runs as part of the design analysis and counts documents by their creation and modification dates (see the table above). The full content analysis is a separate scan, described in this chapter.

An iDNA user starts a content scan for one database at a time. Content analysis is licensed as a number of replica sets. Only the first scan of a replica set counts against the license. Scanning it again, or scanning another replica of it, does not count.

The scan reads the metadata of every document: form name, dates, modifiers, signed and encrypted flags, the names of encryption keys, whether reader and author fields are present, attachment names, sizes and dates, document links, and counts of rich-text elements. iDNA stores aggregates of these, such as attachments by type and size, and modifiers with their counts. The scan reads no item values and no attachment contents.

Is Document or Mail Content Read?

- **Application documents:** metadata only, except Profile Documents, which are read in full.
- **Mail files:** listed in the database inventory and seen in the session records. Design analysis, Profile Documents and content age do not open them, because mail files are not Focus Databases.
- **Mail messages:** Mail Routing Activity, which is not enabled by default and is licensed separately, reads the router's log entries and takes sender, recipients, size and time from them. **It reads no subjects and no message bodies.**
- **Rich text and attachments:** not read, with two exceptions: DLAU documents are read in full, and content analysis reads attachment names, sizes and dates.
- **Source code:** read and stored in full, including whatever developers wrote into it.

Personal Data

Personal data	Read from
Notes user names	Session records, web logs, events, the user activity records of databases, ACLs, design signers and modifiers, Profile Documents, DLAU and MarvelClient records
Data about people and groups	Person documents and group documents in the Domino Directory
IP addresses and client details	Web logs (client address, host name, user agent, requested URL); MarvelClient records (IP address, operating system user name)
Mail addresses	The router's log (Mail Routing Activity); the Domino Directory
Values written by applications or developers	Profile Documents, event messages, source code

iDNA stores all collected data on the appliance with real names. By default, the analyses show user names pseudonymized. iDNA Applications Data Flow in Detail describes which analyses show real names on purpose, who sees what, and what the AI components can reach.

On a Trial License

A trial license reads from the same sources, with the same access, as a full license. This includes the content packages and the design of every Focus Database and template, which iDNA exports and keeps as described above. The differences:

- **Source code:** depending on license options, source code from a subset of databases is analyzed.
- **Profile Documents and content age** are read only for the trial sample.
- **The trial sample** is a set of up to 50 Focus Database instances that iDNA selects itself from the middle of the pack in usage and complexity. iDNA recomputes it every night, so it can change while usage data builds up. You cannot choose it.
- **Content analysis** works as on a full license.

What a trial shows, and where it masks names, is described on iDNA Applications Trial License.

Related Topics

- iDNA Applications Data Flow in Detail: storage, processing, who sees what, what leaves your network, AI clients.
- System Requirements: the access and server settings iDNA needs.
- Admin Interface and Config Manager: where the collection is set up.
- Content Age and Data Retention Analysis in iDNA Applications: what content age measures.
- iDNA Applications License Types and Options: license types and options.
- iDNA Applications Trial License: everything about a trial.
- Database Recategorization: adjusting which databases are Focus Databases.
- Configuring Profile Document Collection in iDNA Applications: setting up Profile Document collection.
- Content Packages - iDNA Applications: setting up the content packages.