Techdio ApS

Data Quality App for Dynamics 365: User Guide

Revised 2026-11-20. Online version: https://www.techdio.dk/learn/data-quality-app-admin-guide/

Data Quality App user guide

User guide for the Data Quality App by Techdio ApS, for Microsoft Dynamics 365 and Dataverse. Release 2026.09.03, revised 2026-11-20. Download the PDF version.

About this guide

This guide takes you from an empty organisation to a working set of duplicate rules, match keys, data quality rules and a dashboard, in the order you would do it. Each chapter starts with what the feature is for, then how to do it, then what to watch out for. Page names are written in bold, as they appear in the app's left navigation: Duplicate Rules, Match Keys and so on.

The ideas in five minutes

Idea	What it means here
Duplicate rule	A definition of "these two records are the same thing": a table, one or more field comparisons, and what happens on a match. Rules run when a record is saved and in bulk runs over a whole table or view.
Field comparison	One row of a rule: compare this column with that column, this way (exactly, first letters, contains, two words in common, similar enough). Rows in one alternative must all match.
Alternative	A group of comparisons. A rule may have several; a pair is a duplicate if any alternative matches. "Same email, OR same phone and surname" is a rule with two alternatives.
Match score	Each comparison carries a weight; a matching pair scores the sum of its weights, adjusted for how rare the matched value is and for other alternatives it also satisfies. Two bars on the rule: pairs above the high bar are confident duplicates, pairs above the review bar are worth a look.
Match key	A second column the product computes from a field and keeps beside it, so rules can compare on something the person did not type: how a name sounds, a phone number in one shape, a company name without its legal form. Rules match on key columns with an ordinary exact comparison.
Match synonym	A word your organisation teaches the matching: an abbreviation, a nickname, a legal form, a word to ignore. Laid over the words the product already knows.

Run	One bulk detection over a table or view, with its results grouped per record. Reviewed and merged from Duplicate Results.
Merge	Two records become one: related records move to the survivor, chosen field values are kept, the loser is deactivated and marked as merged. There is no undo.
Not a duplicate	A pair somebody has ruled out. It is never offered again, on save, in a run or by a merge, until the decision is removed.
Data quality rule	A per-field check on a record, built from one of 22 rule types (a required field, a format, an allowed value, a date, recent activity and more). Each rule passes, warns or fails, and together they give the record a score from 0 to 100.
Job	Bulk work in the background: a detection run, a bulk merge, a match key fill, a score recalculation. Jobs can be scheduled and can repeat.

Installing and setting up

1. Install the app from Microsoft Marketplace

The Data Quality App is installed from Microsoft Marketplace straight into your Dynamics 365 or Dataverse environment. Nothing is downloaded and no file is imported by hand. You need an account that is a System Administrator in the target environment.

1. Open the Data Quality App listing on Microsoft Marketplace and choose Get it now. If you are not signed in, sign in with the work account that administers your environment.
2. Microsoft shows a short form with your name, email and company. Check the details, accept the terms and choose Continue.
3. You are taken to the Power Platform admin centre, to a page named Install Data Quality App. In Select an environment, pick the environment the app should be installed in. Tick the two agreement boxes, the terms of service and the privacy statement, and choose Install.
4. The installation takes a few minutes. You can follow it in the admin centre under Environments, your environment, Dynamics 365 apps, where the status changes from Installing to Installed. The install adds the Data Quality App, its tables, plug-ins, pages and four security roles: Duplicate Detection, Duplicate Detection Bulk Operations, Duplicate Detection Administrator and DQA MCP (with Read, Create and Delete levels).
5. Open the Data Quality App from the app list at make.powerapps.com or home.dynamics.com. The License page shows that the 15-day trial has started with all features, one active duplicate rule and at most ten merges per operation.
6. Repeat step 1 for every additional environment, for example a test environment before production. Each environment gets its own trial.

Installing from a solution file instead. Organisations that receive the solution from Techdio as a file, for example for an offline environment, import it in the Power Platform admin centre: open the environment, choose Solutions, Import, and pick `DataQualityApp_<version>_managed.zip`. The result is the same as the Marketplace install.

2. Give people a role

Assign roles in the Power Platform admin centre. Changes take effect within a minute.

Role	Who gets it
Duplicate Detection	Everybody who works with records: sees duplicate warnings on save, reviews results, merges within their own privileges, marks pairs as not duplicates.
Duplicate Detection Bulk Operations	The people who run bulk detections and bulk merges across a table or view.
Duplicate Detection Administrator	The people who define rules, match keys and synonyms, run bulk detections, schedule jobs and see licence warnings.
DQA MCP (Read, Create, Delete)	Only for people who use an AI assistant with the app (chapter on the assistant). Granular: Delete does not include Create.

3. Start the background engine

Bulk work runs in the background, in one of three ways, and only one per environment:

- **The heartbeat flow**: a Power Automate flow, no extra hosting, quickest to start.
- **An Azure WebJob**: runs in your own Azure, best for large volumes.
- **A Windows service** on your own server.

To start with the heartbeat flow: the solution ships the flow DD Background Engine Heartbeat, switched off.

1. In the environment's solutions, open the Data Quality App solution, find the flow's **connection reference** for Microsoft Dataverse and map it to a connection owned by a dedicated service account or application user, not a personal login.
2. Turn the flow **on**.
3. Start a small detection from a rule and confirm it begins within about a minute. That proves the connection works.

The other two options. Organisations with large volumes, or that prefer their own server, run the same engine as an Azure WebJob in their own Azure or as a Windows service. The next chapter, Installing the background engine, walks through both.
If nothing runs. Jobs that stay "Scheduled" mean no engine is processing them: the flow is off, the connection is broken, or the service is stopped. Scheduled Jobs shows the queue.

4. Licence

A new organisation starts a 15-day trial with full functionality and two limits: one active duplicate rule and at most ten merges per operation. After purchase, Techdio activates the licence for your environment automatically: there is nothing to enter, and the License page shows the new status. For environments that must never contact Techdio, choose offline activation: request the activation code by email from support@techdio.dk, paste it on the License page and press Activate now. Only administrators see expiry warnings; end users are never nagged.

Installing the background engine

Bulk work, detection runs, bulk merges, match key fills and data quality calculations, is done by a background engine that polls the environment for scheduled jobs. There are three ways to run it: the **heartbeat flow** from the previous chapter (a Power Automate flow, no extra hosting, quickest to start), an **Azure WebJob** (runs in your own Azure, best for large volumes) or a **Windows service** on your own server. The WebJob and the Windows service, described in this chapter, are the same program packaged two ways. Choose one of the three; only one engine may serve an environment at a time.

Engine	Choose it when	What it needs
Heartbeat flow	You want no extra hosting and the quickest start.	A Dataverse connection mapped on the flow. Runs under Power Automate's limits.
Azure WebJob	You have large volumes: large tables, detections over hundreds of thousands of records, or work that should run outside Power Automate's limits. Runs in your own Azure.	An Azure subscription and an application user in Dataverse. Techdio's deployment template creates the Web App for you.
Windows service	You want it on your own server, on-premises or a VM.	A Windows machine that stays on, .NET Framework 4.6.2 or later, and an application user in Dataverse.

Keep the engine and the solution at the same version. The engine package and the solution ship together and must be updated together: an older engine against a newer solution, or the other way round, can leave jobs failing at their first step. The Downloads section on techdio.dk/learn always offers the packages that match the current release.

1. Create an application user for the engine

The engine signs in as an application user with a client secret, not as a person. This is done once and serves both the WebJob and the Windows service.

1. In the Microsoft Entra admin centre, register an application (App registrations, New registration, single tenant). Note its **Application (client) ID**.
2. Under Certificates & secrets, add a client secret. Copy the secret value now; it is shown only once. Note its expiry date and put a reminder in your calendar: when it expires the engine stops signing in.
3. Leave API permissions as they are. The engine signs in as the application itself, not on behalf of a person, so the Dynamics CRM **user_impersonation** permission is not needed and adding it changes nothing. Access comes from the security roles of the application user in the next steps.
4. In the Power Platform admin centre open the environment, choose Settings, Users + permissions, Application users, New app user, and pick the registration.
5. Give the application user the Duplicate Detection Administrator role, and the roles that let it read, write and merge the tables your rules cover. A merge deactivates the losing record and moves its related records, so the user needs at least the privileges a person merging by hand would need. System Administrator is the simplest choice where policy allows it.

6. Build the connection string. The URL is your environment's, the ID and secret are from the registration:
   `AuthType=ClientSecret;Url=https://myorg.crm4.dynamics.com/;ClientId={application id};ClientSecret={secret};`

Do not use a person's login. A user name and password in the connection string stops working at the next password change or multi-factor prompt, and every merge would carry that person's name in the audit fields.

2. Option A: the Azure WebJob

There are two ways to set up the WebJob. The quick way uses a deployment template prepared by Techdio: you fill in three values and Azure creates everything. The manual way, further down, goes through the Azure portal step by step and is useful when your organisation wants to create the resources itself.

The quick way: deploy the template

You need the three values from step 1: the environment URL, the Application (client) ID and the client secret. Everything else is filled in.

1. Open the Techdio engine deployment template. It opens the Azure portal on a page named Custom deployment, with the template already loaded. Sign in with an account that may create resources in your Azure subscription.
2. Under Project details, choose the **Subscription** and a **Resource group**. Create a new one, for example `dqa-engine`, so the engine's resources stay together. Choose the **Region** closest to your Dataverse environment.
3. Fill in the three values: **Org Url**, your environment URL such as `https://yourorg.crm4.dynamics.com`; **Client Id**, the Application (client) ID; **Client Secret**, the secret value.
4. Leave the remaining fields at their defaults unless Techdio support advises otherwise: **App Name** (a unique name is generated), **Plan Sku** (B2, a sensible start; S1 or above for very large tables), **Sleep Time Seconds** (15), **Detection Parallelism** (8) and **Package Url** (the engine package published by Techdio).
5. Choose Review + create, then Create. The deployment takes three to five minutes. It creates the App Service plan and the Web App, switches Always On on, stores the settings and the `CRM` connection string, and uploads the engine as a continuous WebJob on a single instance.
6. When the page says the deployment is complete, open Outputs. **webApp** links to the Web App and **webJobLogs** to the engine's log. The first lines of the log show the engine version, the settings and the user it signed in as.
7. In the Data Quality App, start a small detection from a rule and confirm under Scheduled Jobs that it moves to In Progress within about a minute.

Updating with the template. When a new release ships, open the same template link again, choose the same resource group and enter the same three values. The deployment replaces the WebJob with the current engine package and keeps the Web App and its settings.

Cost. The template creates a Basic B2 App Service plan by default. Azure bills the plan by the hour while it exists; delete the resource group to stop the cost. Check the current price for your region in the Azure portal before creating it.

The manual way: create the Web App yourself

The package is `DuplicateDetection.BackgroundService.Standalone.zip`, available under Downloads on techdio.dk/learn. It is uploaded as it is; do not unzip it. A step-by-step walkthrough with screenshots of the Azure portal is in Azure WebJob setup for the Data Quality App background service.

1. In the Azure portal create a **Web App**. Runtime stack: **ASP.NET V4.8** (.NET Framework, not .NET Core). Operating system: Windows. Pick a plan meant for continuous work, such as B2 or S1: the engine runs several worker threads, and the free and shared tiers throttle them.
2. Open the Web App, then Configuration, General settings: switch **Always On** to On and save. Without it the site unloads after a while and the engine stops polling.
3. Configuration, Connection strings: add one named exactly `CRM`, type **Custom**, value the connection string from step 1. Save.
4. WebJobs, Add: name it, upload the zip, type **Continuous**, scale **Single Instance**. Single instance matters: two copies would compete for the same jobs.
5. Refresh until the job shows **Running**, then open its Logs. The first lines show the engine version, the user it signed in as, and every effective setting.
6. In the Data Quality App, start a small detection from a rule and confirm under Scheduled Jobs that it moves to In Progress within about a minute.

Settings. All optional, entered under Configuration, Application settings; an unconfigured WebJob already runs at the recommended values. Changing a setting restarts the job, which is safe: an interrupted job resumes from its saved state. The most useful ones:

Setting	Default	Meaning
SleepTimeSeconds	15	Seconds between polls when there is nothing to do.
BulkDetectionBatchSize	200	Records checked per background step.
BulkDetectionDegreeOfParallelism	8	Concurrent detection workers. Lower it on a small plan.
BulkMergeMergesPerStep	20	Merges per step before progress is saved.

The remaining settings, which switch the engine between running its work inside Dataverse or from the WebJob, are listed in the README inside the package. Leave them at their defaults unless Techdio support asks otherwise.

Updating. When a new solution version is imported, upload the matching package: open the WebJob, stop it, delete it, and add the new zip the same way. The connection string and settings stay on the Web App. If you deployed with the template, run the template again instead.

3. Option B: the Windows service

The package is `DuplicateDetection.BackgroundService.Installer.zip`, available under Downloads on techdio.dk/learn. Unzip it on the machine that will run the engine; it holds `setup.exe`, the installer file and a README.

1. The machine needs .NET Framework 4.6.2 or later; Windows 10, Windows 11 and Windows Server 2016 or later include one. It must stay on: a laptop that sleeps is not an engine.
2. Run `setup.exe` and follow the wizard. It asks for administrator rights itself.
3. When prompted, paste the connection string from step 1. It is encrypted for this machine and stored under `%ProgramData%\Cowia\DuplicateDetection`, not in any configuration file. The same prompt can be run again later from the installation folder to change it.
4. The wizard starts the service. If it did not, open Services and start **DD Background Service**. In Add/Remove Programs the same thing is listed as Duplicate Detection Background Processing Service.
5. Check the `logs` folder in the installation directory: the first lines show the version, the user the service signed in as, and its settings. Then start a small detection in the app and watch it move to In Progress under Scheduled Jobs.

Settings are the same as for the WebJob, placed in `DuplicateDetection.BackgroundService.exe.config` in the installation folder, where each one is already present, commented out and explained. Restart the service after a change.

Updating. Run the new version's `setup.exe`. It offers **Update**, which removes the installed version and installs the new one while keeping the saved connection, and **Remove**. Uninstalling through Add/Remove Programs leaves the stored connection behind, which can be deleted by hand.

4. If jobs stay Scheduled

What the log says	What it means
No CRM connection string found	On Azure the connection string is not named `CRM` or is not of type Custom. On a machine the stored connection is missing: run the connection prompt again.
Could not connect to CRM with the configured connection string	The string was found and rejected: wrong URL, expired secret, disabled application user, or a user without privileges.
Signed in, but nothing happens	No job is Scheduled, or two engines are running against the environment. Check the flow is off if you installed a WebJob or service.
Jobs run but slowly	Throughput is bounded by Dataverse's service protection limits, about 6,000 requests per five minutes per user. A bigger plan or machine does not lift that; the parallelism setting past the point where Dataverse throttles gains nothing.

A tour of the app

Navigation	Page	What you do there

Overview	Data Health Studio	The page the app opens on. Your data quality dashboard: duplicates found and merged, rules passing and failing, over time. Arrange it as you like; the layout follows you.
Duplicates	Duplicate Rules	Every duplicate rule and its health. Create, open, estimate, run.
	Duplicate Results	Every detection run and what remains to review. Open a run to review and merge.
	Not Duplicates	Pairs somebody decided are not duplicates. Remove a pair to undo the decision.
Data Quality	Data Quality Rules	Every data quality rule and its health. Create and open rules.
	Data Quality Score	Per table: which rules count, how warnings weigh, whether scores recalculate automatically; recalculate.
Administration	Scheduled Jobs	Every background job and its status. Cancel, retry, start early, open the related rule or run.
	Match Keys	Every match key and its health. Create, fill, audit, rewrite.
	Match Synonyms	The words your organisation teaches the matching, and the words the product already knows.
	Translations	The app’s own texts, translated or reworded. Changes the interface, not your data.
	License	Status, expiry, days remaining, activation.

Duplicate rules

What a rule is for

A rule says when two records of a table are the same thing. It runs in two places: when a user saves a record, the rule checks it against the rest of the table and warns or blocks; and in a bulk run, where every record in a table or view is checked and the pairs are grouped for review.

Creating a rule

Open Duplicate Rules and press New rule. The page asks its questions in order and creates the rule once, when you press Create.

1. **Which table?** The table whose records are checked (the source), and the table they are checked against (the target). Usually the same; a Contact-against-Lead rule is allowed.
2. **Which fields must match?** Add a comparison row per field. For each: the source column, the target column, and how to compare. If the column has a match key, the page offers to match *by how it sounds*, *as normalised text*, *as a telephone number* and so on; that is an exact comparison on the key column. Put rows that must all be true in one alternative; add a second alternative for "or". Weights and the two score bars live here too; the defaults are sensible.
3. **When does it run?** When a record is created or changed, only on create, or only on change.
4. **Who runs it?** The saving user, the system, or a named user. Choose the system or a named user when ordinary users cannot read every record the rule should compare against.
5. **Which records are checked?** Optionally limit the source and the target to a saved view or a FetchXML filter. Only records in the filter take part.
6. **How are duplicates shown?** The result view decides which columns the duplicates dialog and the results pages show. Whether the save is only warned about or blocked. Whether inactive records count.
7. **Name the rule** and create it. Rules are active at once.

The six comparisons

Comparison	Example pair it catches	Notes
Exact	Nordisk Vinduer = nordisk vinduer	Case-insensitive, like the database. The comparison to use on a match key column.
First N characters	Kastrup Logistik A/S ≈ Kastrup Transport (N=6)	A value shorter than N compares whole.
Last N characters	98 44 22 33 ≈ 98 45 22 33 (N=4)	For a phone column a telephone key is usually better: it understands country codes.
Contains	Bornholm Byg ⊂ Bornholm Byg og Anlæg	One direction: the stored value contains the typed one.
2 Word Match	Aalborg Marine Service ≈ Marine Service Aalborg	Two whole words in common, any order. Punctuation glued to a word is not part of it.
Similar enough (N = %)	Christianssen Revison ≈ Christiansen Revision (85)	Typo tolerance. Reads accents both ways and words in fixed order. On a column with a normalised, company or sounds-like key it searches on its own; without one it needs an exact or starts-with row beside it in the same alternative.

Per-row options. Trim, remove all whitespace, compare blank values (two blanks count as equal), compare lookups as text (by the related record's name), ignore words (a comma list removed before comparing, for plain text columns), and a date-time behaviour for date columns. Non-text columns compare by equality.

Start from the starter rules. The import page offers three built-in bundles, Accounts, Contacts and Leads, that set up the match keys and a rule each with sensible alternatives. Most organisations need

nothing else to begin. See the export and import chapter.

What a user sees on save

When a save matches a rule, the form shows *Duplicate found for this record* with View duplicate records. The dialog lists the duplicates in the rule's result view, grouped by the rule that found them, and lets the user open a record or merge. A rule set to block refuses the save until the duplicate is dealt with; a rule set to warn lets the user save anyway.

Health and estimate

Duplicate Rules shows each rule's health: rules that cannot run, have no fields, reference deleted columns or views, use a disabled run-as user, duplicate another rule, can never reach their own review bar, match through a key that was never filled, or rely on a telephone key with no country. Open a rule for an **estimate**: a sampled count of how many duplicates it would find, before running it.

Match keys

What a key is for

People write the same thing differently. A match key is a second column the product computes from a field and keeps up to date, so a rule can compare on the computed value with an ordinary exact match, which is fast and indexed. Adding a key changes no rule; a rule uses a key by matching on its column.

The eight kinds

Key	Example: value → key	Choose it when
Sounds like	Jorgensen → JRJNSN Joergensen → JRJNSN	Names spelled several ways that sound the same. English, German or Spanish. Two readings, so a name that sounds two ways is matched on either.
Normalised text	Vejle Bryghus → BRYGHUS VEJLE Grøn → GRON / GROEN	Any text where case, punctuation, spacing, special letters or word order vary. Can be a composite of several columns.
Company name	Herning Maskinfabrik A/S → HERNING MASKINFABRIK Aarhus Intl Trading Ltd → AARHUS INTERNATIONAL TRADING	Company names. Strips legal forms for the countries you choose, writes out shortenings, recognises a form at either end, ignores word order.
Telephone	+45 66 12 34 56 → +4566123456 0045 66123456 → +4566123456	Any phone column. Give it a country so bare national numbers get their code; or let it read the country from a column of the record.
Email	Info+crm@Techdemo.dk → info@techdemo.dk	Email columns. Lower case, +tags removed, Gmail dots, googlemail = gmail, first address of several.

Registration number	DK 25 31 37 63 → 25313763 (verified)	CVR, VAT and company numbers. Country prefixes stripped, checksums verified for DK, SE, NO, FI, DE, GB, NL and CH; a verified number weighs more.
Web domain	https://www.koldingmoebler.dk/ → koldingmoebler.dk	Website or email columns. The strongest company identifier after a registration number.
Given name	Bob → ROBERT Kristian → CHRISTIAN	First-name columns. Weak alone; pair it with a surname or a postcode in the same alternative.

Creating a key

1. Open Match Keys, press New match key.
2. Choose the table and the field, then the kind. The page proposes a column name; you may edit it.
3. Answer the kind's own questions: the language for a sounds-like key; the country for a telephone or registration key, or the column to read each record's country from; for a company key whether to ignore legal forms and for which countries; for a normalised key whether to ignore word order and which other columns to key together with it (a composite key).
4. Create. The key column and its second-reading column appear on the table, and records saved from now on are keyed at once.
5. Fill existing records: the page offers to queue the fill as a background job. Until it finishes, rules on the key find only records saved since.

A column may carry one key of each kind. A company name can have a sounds-like key, a normalised key and a company key side by side; a "similar enough" comparison searches through all of them.

Country from a column

A telephone or registration key with customers in several countries should read the country from the record. Choose the column on the New Match Key page; text, choice and lookup columns all work, and "Danmark", "Germany" and "DE" are all understood. The fixed country is the fallback for a record with none. Changing a record's country re-keys its number.

Keeping keys healthy

Keys are written by the product when a record is saved. Anything that writes around it, a data import with plug-ins switched off, an integration that bypasses custom plug-ins, leaves keys empty or stale. Match Keys audits each key: records with a value and no key, and a re-keyed sample of the most recently changed records. Rewrite all keys recomputes every record. Changing a key's country, language or options marks it "definition changed" until you rewrite it.

Deleting a key deletes its columns and every value in them, and is refused while any rule still matches on them. Fix or delete those rules first.

Match synonyms

Match Synonyms holds the words your organisation teaches the matching. They are laid over the words the product already knows; press Show the words the product already knows on the page to see those. A row with the same word as a built-in entry wins over it; a row that maps a word to itself switches the built-in entry off.

Kind	Example	Read by
Company abbreviation	Manufact = Manufacturing Co = Co (switch the built-in off)	Company name keys that write out shortenings, and “similar enough” comparisons.
Given name	Kat = Katarina	Given name keys, and “similar enough” comparisons.
Legal form, with country	Kft (HU)	Company name keys set to ignore the legal form for that country. Nothing else: this is the one kind that needs a match key.
Word to ignore	Holding	Normalised text and company name keys, and “similar enough” comparisons. The tool that works where a rule’s “ignore words” does nothing, because keys are computed before rules run.

Which kinds need a match key. An abbreviation, a given name and a word to ignore work without any key, through a rule field set to the “similar enough” comparison: both values are read several ways before scoring, with shortenings written out, a first name replaced by the name it stands for and ignored words dropped, and the best reading counts. A legal form is only stripped inside a company name key set to ignore legal forms for its country, so a legal form saved for a country no key reads is kept but idle until such a key exists; the page says so under the form. Exact, starts-with and sounds-like comparisons never read synonyms.

Adding a word. Pick the kind, type the word as it appears in the data and, where the kind has one, what it stands for. For a legal form, choose the country from the list; the countries your company name keys already strip are listed first, and a country no key reads yet shows a note. The product stores it in upper case without punctuation, names the row, and refuses what makes no sense: a legal form without a country, a given name of two words, the same word twice, or a chain (Mfg means Manufact while Manufact means Manufacturing).

After adding a word: rewrite the keys

A word takes effect in two steps, and the second one is yours to do.

1. **New and changed records are keyed with the word at once.** From the moment the row is saved, every record saved or changed is keyed using it. A company name key with shortenings written out stores “Herning Manufact. A/S” as HERNING MANUFACTURING right away.
2. **Records that already exist keep their old keys until you rewrite them.** Their key columns were computed before the word existed, and nothing touches a stored record because a table row changed. So saving the word marks every match key that reads that kind of word as “definition

changed, rewrite needed”. Open Match Keys, find the marked keys, and press Rewrite all keys. The rewrite runs as a background job and recomputes the key for every record of the table; the job appears under Scheduled Jobs.

Until the rewrite has run, a rule on that key finds the new word only among records saved since the row was added. Which keys are marked depends on the kind of word: an abbreviation or a legal form marks company name keys; a word to ignore marks normalised text and company name keys; a given name marks given name keys. Sounds-like, telephone, email, web domain and registration keys are never marked, because they never read the table. The same two steps apply when you switch a word off or remove it, and when you change a key’s own settings, such as its country or its word-order option.

Add several words, then rewrite once. Each row marks the same keys, so enter the whole batch first and run one rewrite per table, rather than one after every word.

Running detection

On save

Every active rule runs when a record is created or changed, according to its trigger. Nothing to start.

A bulk run

1. Open the rule from Duplicate Rules and choose to run it. Optionally limit the run to a saved view.
2. Choose **now** or a **scheduled time**, and whether the job **repeats** every N minutes, hours, days or months. A repeating job creates its next occurrence when this one finishes.
3. The run appears in Duplicate Results and the job in Scheduled Jobs. Detection is done in batches inside Dataverse; the job records progress and a log.

Auto-merge. A rule can carry auto-merge settings, so a bulk run merges what it finds above the high bar without anyone reviewing, using the survivorship you chose. Use it for rules you trust; the review bar and the high bar are what make it safe.

Reviewing results

Duplicate Results lists every run: rule, table, when, how many records with duplicates, how many remain to review. Filter by table. Open a run.

A run page groups the duplicates per record. Each row shows the record, the number of duplicates and the best match score; pairs at or above the review bar are marked as worth reviewing, pairs at or above the high bar as confident. Click a record to see its duplicates individually, which alternative of the rule joined each pair, and the columns of the rule’s result view side by side.

From here you can Merge duplicates for one group, Merge all duplicates for the run (now, or as a Background merge job), mark a pair as not a duplicate, Open printable report, or Delete this run and its results. Deleting a run deletes only the result rows; the records are untouched and merges already done are not undone.

Merging

A merge combines two records. Related records, both one-to-many and many-to-many, are moved onto the surviving record; the field values you choose are kept; the loser is deactivated and marked as merged, pointing at the survivor.

There is no undo. A merged record can be reactivated, but its related records stay where they were moved. Review before merging, and use the review bar.

Kind of merge	Where	How the survivor is chosen
Single	From a record's duplicate dialog or a result group	You pick the survivor and, field by field, which value to keep.
Group	A group on a run page	All duplicates of the group merge into its anchor record.
Bulk	A whole run, now or as a background job; or auto-merge on the rule	Survivorship per group: oldest, newest, most activities, most filled fields, or best data quality score.

Minimum match score. A bulk merge can be limited to duplicates whose match score reaches a number you set. The box starts at the rule's high bar, the score at which the run page calls a duplicate confident, so ticking it means "merge only what the run page calls confident". Duplicates below the score are skipped and stay open on the run for you to review; so are duplicates without a score.

Options. Include or exclude each relationship from reparenting; skip errors and continue; bypass synchronous plug-ins, asynchronous plug-ins or Power Automate flows during the merge; which inactive status the loser gets. A merge of a record somebody else is merging at the same moment is refused by the platform: wait and run again, and already-merged records are skipped. A trial licence allows ten merges per operation.

Not Duplicates

When two records look alike but are not the same, mark the pair as not a duplicate from the run page or the duplicate dialog. The pair is then never offered again: not on save, not in a bulk run, not by a merge. Not Duplicates lists every such decision with who made it; removing a pair undoes the decision. A user may undo their own; an administrator may undo anybody's.

Export, import and starter rules

Starter rules

Open Duplicate Rules, choose Import. The page offers three bundles made by the product:

Bundle	Keys it sets up	Rule

Accounts	Registration number, web domain, company name, sounds-like on the name	The same company however it was written: by registration number, or web domain, or company name with the postcode, or the sound of the name, or a name spelled nearly the same, or a looser likeness with the postcode.
Contacts	Email, mobile phone, given name, normalised surname	The same person: by email, or mobile and surname, or first name (Bob = Robert) with a similar surname and the postcode.
Leads	Email, business phone, company name, normalised surname	The same lead: by email, or business phone and surname, or company name with a similar surname.

Importing a starter creates the keys that are not already there, queues their fill, and creates the rule. Importing again updates rather than duplicates.

Moving rules between environments

1. In the source environment, select rules on Duplicate Rules and export. The zip carries the rules, the match keys they match through and the synonyms those keys read, with a manifest and notes.
2. In the destination, open Import and choose the zip. Nothing is written yet: the page says what would happen to each rule and key, asks you to confirm a person or view it matched by name, and blocks a rule that cannot work here.
3. Tick what to import and import. Keys are created first with their columns; a key of the same kind already here is used instead, and the rule's fields point at its column.

Formats. A bundle exported by a newer release than the destination reads is refused in words rather than imported wrongly. Update the destination first.

Data quality rules

A data quality rule checks one thing about a record and gives a verdict: pass, warning or error. Create them on Data Quality Rules; the rule builder asks for the table, the rule type, the field and its parameters, the severity, and the scope. Rules are unlimited, also during the trial. The app has 22 rule types:

Group	Rule types
Completeness	Required; At least one of these fields (for example email or phone or mobile); Minimum length; Must be empty.
Format and validity	Format (email, phone, web address, post code); Regular expression; Maximum length; No stray spaces; Must not contain (filler such as “test” or “n/a”).
Allowed values	Must be one of; Must not be one of; Number range; Number of selections (multi-select choices).
Dates and consistency	Date check (past, future, within or older than N days); Date order; Must match another field; Must differ from another field; Linked record still active.
Freshness and engagement	Last modified; Last activity; Number of activities (filtered by activity type, state and direction).

Your own logic	Power Automate flow: a flow decides, for checks that need an outside source or business logic.

Preview before you save. The rule builder shows how many existing records the rule would fail, with examples, and lets you test your own values. Nothing is saved until you confirm.

Severity. A **warning** counts as half a pass in the score by default (you can change the weight). An **error** counts fully and can, if you switch it on and confirm, block saving on create, on update or both. **Scope**: a rule can apply only when a field has a value, only to a business unit or team, or only to records in an additional filter, and can ignore chosen users such as integration accounts. **Weight**: a rule can count double or be flag only, which means it is excluded from the score while still flagging records. New and changed records are evaluated on save; existing records need a recalculation job.

Export and import. Rules can be exported as a bundle and imported into another environment, for example from development to test to production (see Export, import and starter rules).

Data Quality Rules shows each rule's health: rules that cannot run, reference deleted fields or duplicate another rule.

Data quality score

Each record gets a score from 0 to 100: passes plus half the warnings, divided by the rules that apply, times 100. Rules that do not apply to the record are left out. On the record form a dashboard shows each rule's outcome.

Data Quality Score configures each table: which rules are included, how warnings weigh, whether scores recalculate automatically, and a button to recalculate the whole table as a background job. Recalculate after you change rules or weights, so existing records catch up.

Data Quality 360

Data Quality 360 combines the data quality of a record and its related records into one score, so you see the whole customer and not only one record. Open it with Data Quality 360 in the top right corner of the Data Quality Dashboard on any record with data quality rules. Nothing extra is installed or configured: it uses the rules you already have.

- **Reach.** This record scores the record alone; + Related adds its directly related records; + 2 steps and + 3 steps go one and two steps further. Each record is counted once, and the combined score is weighted by how many records there are.
- **Start anywhere.** Start from an account, a contact, a case or any table with rules. From a contact it goes up to the account and from there to the account's other related records.
- **Tables.** Tick the related tables that should count in the combined score.
- **What to fix first.** The failing rules across all the records are ranked, errors first, then by how many records fail them. Show them filters the view to exactly those records.
- **Export.** Export an Excel workbook with every record the score reaches, the rules each one fails and a link to it, or print the view as a PDF. The file is built in your browser.

Rules marked as flag only are shown but do not change the score. Because Data Quality 360 uses the same rules, its numbers match the Data Quality Dashboard on each record and the trends in Data Health Studio.

Data Health Studio

Your data quality dashboard. Widgets show open duplicate pairs, records merged, failed merges, active quality rules, the data health score (90 and above is excellent, below 40 needs attention), duplicates found by month, and merge status. The header applies to every widget at once: date range (30 days, 90 days, this year, a custom period or all time), record types, whole organisation or my records, refresh.

Press Edit to arrange it: add, remove, resize and move widgets, and change what each shows. Your layout is saved to your user and follows you; an administrator can publish a default for everyone.

Scheduled jobs

Scheduled Jobs lists every background job: bulk detection, bulk merge, match key fills, data quality calculations and health snapshots. Each shows its status and what it is working on, and keeps a newest-first log.

Status	Meaning and what you can do
Scheduled	Waiting for its time, or for the engine. Start it early, or cancel it. If everything stays Scheduled, the engine is not running (see installing).
In Progress	Working. Cancel stops it at its next step and keeps the work done so far.
Done	Finished. A repeating job has already created its next occurrence.
Error	Stopped on a failure; the log says where. Retry runs it again as a fresh job.
Canceled	Stopped by a person.

Translations

Translations lists every text the app shows its users: duplicate checks, merge screens, dashboards. Edit the wording or translate it into another language. This changes the interface only, never your data. Texts are grouped by screen; an entry with no translation for a language falls back to English.

License

License shows the licence type, its expiry, days remaining, and the trial limits while on trial: one active duplicate rule and ten merges per operation. After purchase the licence is activated automatically; there is nothing to enter. For offline activation, paste the activation code you received from support@techdio.dk and press Activate now. After expiry there is a grace period; then new rules and bulk operations are

refused until a licence is applied, while existing real-time checks keep running. Warnings are shown to administrators only.

The AI assistant (MCP)

The DQA MCP gateway lets Copilot and other AI agents read data quality, create rules and run detections, and merge only with your approval, gated by security roles. It connects an AI assistant, such as Claude, Copilot Studio or ChatGPT, to your environment, so you can manage duplicates and data quality by talking to it. The server is software you run, on your own machine or in your own Azure tenant; your data flows only between you and your Dataverse environment.

1. Unzip `DqaMcp-win-x64-<version>.zip` (or the Linux build for a server) and register `dqa-mcp.exe` in your assistant. No configuration is needed: the first time you ask it something, a browser window asks for your normal Microsoft sign-in and the assistant lists the environments you can access.
2. Have an administrator assign you one of the **DQA MCP** roles. Any role lets the assistant read; Create is needed to create rules and run detections; Delete to delete configuration and to merge. Nobody is exempt, System Administrators included.
3. Ask. “How healthy is my CRM data?”, “Check Contoso Ltd for duplicates”, “Set up duplicate detection for leads matching email exactly and last name by sound”, “Add a match key on contact last name and backfill it”, “Teach the matching that Manufact means Manufacturing”, “What background jobs are running?”

Everything the assistant does runs as you, so your Dataverse security applies and audit fields carry your name. Merge tools are hidden unless the machine enables them, and every merge needs your explicit approval. The assistant explains the product from its own built-in documentation and points you to the app page for the few things it cannot do itself: score weights and thresholds on duplicate rules, auto-merge configuration, bundles, licence activation, translations.

Housekeeping and troubleshooting

Symptom	What to check
A run or fill stays “Scheduled”	The background engine: is the heartbeat flow on and its connection valid, or is the WebJob or service running (see Installing the background engine)? Scheduled Jobs shows the queue.
A rule on a match key finds only new records	The key was never filled, or was filled before an import bypassed the plug-ins. Match Keys shows “records without a key” and “keys out of date”; rewrite the key.
Two obvious duplicates are not found	Open the rule’s health and estimate. Common causes: the value differs inside the first letters and no sounds-like key is on the column; the pair is on the Not Duplicates list; one record is inactive and the rule excludes inactive records; the rule’s filter view excludes one of them.
Too many pairs are offered	Raise the rule’s review bar, or make the loosest alternative require a second field such as the postcode.

A word the matching should know	Add it under Match Synonyms and rewrite the keys that read it.
A new column does not appear right away	Dataverse publishes new columns to its query engine and its data API a little after creating them. Wait a minute and refresh.
The page says to open it from inside the app	The app's pages need the app context. Open them from the Data Quality App's navigation, not from a bookmark to the file.
Merges fail for some records	Open the job's log. A record merged by somebody else at the same time is skipped; a relationship the merging user cannot write is reported per record when "skip errors" is on.
After an upgrade an existing rule does not use a new matching feature	The rules' cached definition is rebuilt when a rule or key is saved. Open and save one rule.

Glossary

Term	Meaning
Alternative	A group of comparisons in a rule that must all match; several alternatives are ORed. Also called a match group.
Alternate column	A key's second reading: the other way a name sounds, the expanded spelling of Ø, the dots-removed Gmail form. Rules match on both.
Anchor	The record a result group is built around; the survivor of a group merge.
Backfill / fill	The background job that computes a match key for records that existed before the key.
Composite key	A normalised key computed from a field together with other columns of the record.
High bar	The score at which a pair is a confident duplicate; what an unattended merge must clear. The rule page calls it the strong match threshold; Automatic means 85.
Match key	A computed column beside a field, so rules can compare on it. Also called a phonetic field in older releases.
Match synonym	A word the organisation teaches the matching, laid over the built-in word lists.
Review bar	The minimum score for a pair to be offered for review. The rule page calls it the review threshold; Automatic means 40.
Run	One bulk detection with its grouped results.
Run-as	Whose privileges a rule uses when it searches: the saving user, the system, or a named user.
Rewrite all keys	Recompute a key for every record, not only the ones without a key.
Survivorship	The rule for choosing which record survives a bulk merge.

Data Quality App by Techdio ApS. User guide for release 2026.09.03. Revised 2026-11-20.