



INTUITIONLABS / RESEARCH & PRACTICAL GUIDANCE

FDA DECRS Supplier Verification: Registration Data Pipeline

Build practical AI for pharma and biotech with IntuitionLabs. We help life-science teams turn complex information and workflows into useful software, governed knowledge systems and AI tools.

Published by IntuitionLabs · 2026-10-27 · fda decrs · drug establishment registration · supplier verification · supplier qualification · fda establishment identifier · ndc directory · record linkage · pharmaceutical due diligence

Original article: <https://intuitionlabs.ai/articles/fda-decrs-supplier-verification-data-pipeline>



In this article

1. Executive Summary
 2. Introduction and Background
 3. What DECERS Can and Cannot Answer
 4. Registration and Listing Lifecycle
 5. Current Versus Excluded Records
 6. Entity and Address Normalization
 7. Match Rules: Exact, Fuzzy and Unresolved
 8. Cross-Checks With NDC, SPL, Inspection and Foreign Registers
 9. Discrepancy Triage
 10. Refresh Architecture, Audit Trail and Schema
 11. Case Study: Worked Supplier-Verification Example (Hypothetical Example)
 12. Data Analysis and Evidence
 13. Implications and Future Directions
 14. Frequently Asked Questions (FAQs)
 15. Conclusion
-

Executive Summary

The FDA Drug Establishments Current Registration Site (DECERS) is a public list of "currently registered establishments which manufacture, prepare, propagate, compound or process drugs" (www.fda.gov). FDA states that the file is updated each business day (www.fda.gov), and its search page showed "Data Current through" Monday, October 26, 2026 when retrieved on October 27, 2026, a source age of 1 day (www.accessdata.fda.gov). For supplier-quality and procurement teams, DECERS answers one narrow question well: does FDA's registration system currently list an establishment that matches the entity and address a supplier has claimed? It does not say whether the supplier is approved, inspected, compliant or the right source for a given material. Registration "does not denote approval of the establishment, the drug, or other drugs of the establishment" (www.ecfr.gov).

A reproducible verification pipeline rests on five design choices. First, take dated snapshots of both the current and the excluded download files that FDA links on the DECERS page (www.fda.gov), and hash each file (csrc.nist.gov). Second, normalize names and addresses with versioned rules built on public standards for Unicode, postal suffixes and country codes. Third, grade every candidate with a deterministic rubric and never call a fuzzy match verified. Fourth, compare registration with listing data in the National Drug Code (NDC) Directory (www.fda.gov) and with FDA inspection data keyed on the FDA Establishment Identifier (FEI) (datadashboard.fda.gov). Fifth, send each discrepancy to a human review queue that carries a source date and URL.

Dates drive that queue. Establishments must renew between October 1 and December 31 each year (www.govinfo.gov). A registration received in that window is current through December 31 of the following year (www.fda.gov), and a submission outside it does not extend expiry beyond the current calendar year. As of the October 27, 2026 publication date, a registration expiring December 31, 2026 is 65 days from expiry. Scope

matters as much as dates: DECERS does not contain outsourcing facilities (unless they also registered under other business operations) or wholesale distributors and third-party logistics providers (www.fda.gov), and makers of harmless inactive ingredients such as excipients are exempt from registration.

At population scale, FDA's statutory report counted 13,711 registered drug establishments at the end of fiscal year 2025, of which 4,816 (about 35%) were foreign (www.fda.gov). Counts differ by definition, so any published match rate must therefore state its snapshot, denominator, rules and manual-review exclusions. This report publishes no match rate, because it did not ingest the files. It supplies the method, schema and rubric a team needs to produce its own.

Introduction and Background

Supplier qualification starts with a claim: this company makes this drug or active pharmaceutical ingredient (API) at this address. Before an audit is scheduled, a team can test part of that claim against public regulatory data, a test that is narrow, cheap and testable. If a registration record is absent, stale or attached to a different entity, the claim needs an explanation. If a record is present, the claim still needs primary documents, because registration is a notice to FDA and not a certificate from it. FDA's electronic registration guidance states that registering and listing does not indicate the products are approved or that FDA has verified the information provided.

This report is written for supplier-quality leaders, procurement teams, regulatory operations staff and contract development and manufacturing organization (CDMO) or API due-diligence analysts. Its angle is a pipeline view: DECERS is treated as a dated data source with a lifecycle, a scope boundary and a failure mode for every join. The search form requires a firm name (www.accessdata.fda.gov), which suits a spot check; the bulk files support snapshots, differences and audit trails.

The publisher, IntuitionLabs, is an adjacent advisor to life-sciences organizations and not a data vendor or a party to any supplier list. Nothing here replaces a quality agreement, an audit or legal advice; teams with questions about a specific supplier relationship should consult qualified regulatory counsel.

Three terms are easy to conflate. A **registrant** is "any person that owns or operates an establishment" that manufactures, repacks, relabels or salvages a drug (www.ecfr.gov). An **establishment** is "a place of business under one management at one general physical location" (www.ecfr.gov). A **listing** is the separate duty to list each drug the registrant manufactures, repacks, relabels or salvages (www.ecfr.gov). Legal entity, site and product sit at three levels, and most discrepancies in a verification exercise are mismatches between those levels and not data errors.

What DECERS Can and Cannot Answer

What it can answer

DECERS supports four dated statements: a record matching a given site was or was not present; it carried a particular expiry; it listed particular business operations, such as API manufacture, which FDA's vocabulary distinguishes from other operations; and it was associated with identifiers (DUNS and FEI, below) that can be joined to other datasets.

What it cannot answer

DECRS cannot say whether a supplier is approved, compliant, inspected or authorized to sell a product. Registration does not mean "that a product may be legally marketed" (www.ecfr.gov). It cannot say whether the party quoting a price is the registrant, because a private label distributor that does not itself manufacture, repack, relabel or salvage need not register (www.ecfr.gov). And it omits facility classes that FDA names. The page says the database does not contain human drug compounding outsourcing facilities, although one "may appear in the database if it also registered under other business operations" (www.fda.gov), and it does not contain wholesale drug distributors and third-party logistics providers.

Part 207 adds further exclusions. Excipient makers are exempt (www.ecfr.gov), as are research-only drugs (www.ecfr.gov), and the part does not apply to establishments that solely make medical devices (www.ecfr.gov) or to blood collection and processing establishments (www.ecfr.gov). Absence from DECRS can be legitimate for an excipient maker, yet it still needs an explanation on file.

Identifiers: DUNS, FEI and the unique facility identifier

Federal law requires each registration to carry a unique facility identifier of each establishment (www.govinfo.gov), and the regulation repeats it (www.ecfr.gov). FDA's guidance names the Dun and Bradstreet DUNS number as its preferred identifier. The DUNS must be that "of the establishment (not one linked to the corporate headquarters)" (www.fda.gov), and FDA's training material says "The DUNS Number is site specific" (www.fda.gov). D&B itself says headquarters, branches and subsidiaries "can all be uniquely identified and tracked" (www.dnb.com). The FEI is separate: it is issued by FDA to track inspections, and "An FEI number is a facility specific identifier" (www.fda.gov). FEI strings are 7 or 10 digits, with leading zeros where necessary (www.fda.gov), a concrete reason to load identifiers as text. The Legal Entity Identifier, a 20-character code (www.gleif.org) that "can represent only one entity" (www.gleif.org), names a company and not a plant.

One documentation gap matters. The DECRS page does not say which identifiers appear in the search results or the download files and publishes no column dictionary, and no FDA page found during research listed the zip files' columns. A pipeline should read the header row at ingest, record it, and fail a test when it changes. Supplier master data may also use another identifier family: the U.S. General Services Administration stopped using DUNS in its systems on April 4, 2022 (www.gsa.gov).

Registration and Listing Lifecycle

Initial registration and listing

FDA states that domestic and foreign establishments that manufacture, repack or re-label drug products must register, and registrants must list their commercially marketed drugs. Domestic establishments register "no later than 5 calendar days after beginning to manufacture, repack, relabel, or salvage" (www.ecfr.gov). Foreign establishments register before a drug is imported or offered for import (www.ecfr.gov) and must name a single U.S. agent (www.ecfr.gov). A 2022 amendment extends the foreign duty "regardless of whether the drug or device undergoes" further manufacture elsewhere before import (www.govinfo.gov), and a drug made in an establishment "not duly registered" is deemed misbranded (www.govinfo.gov).

Listing follows. The regulation requires listing information within 3 calendar days after initial registration (www.ecfr.gov) and a review each June and December (www.ecfr.gov). An update is due within 30 calendar days of events such as a name or address change (www.ecfr.gov). A legitimate change can therefore lag a supplier's letterhead by up to a month, so a one-field mismatch is a review item and not a negative finding.

Renewal window and expiry arithmetic

Annual review runs "during the period beginning on October 1 and ending December 31 of each calendar year" (www.ecfr.gov), and the statute uses the same window (www.govinfo.gov). FDA's registration instructions give the rule: establishments renew between October 1 and December 31, and a listing "will show an expiration date of December 31 of the following year" (www.fda.gov).

A pipeline can compute proximity from explicit dates only. With a data date of October 26, 2026 and retrieval on October 27, 2026, days since source refresh = 1. For a record expiring 2026-12-31, days to expiry as of 2026-10-27 = 65 (4 days left in October, 30 in November, 31 in December). The renewal window closes on that same date, so every record expiring December 31, 2026 is inside its open window, and the queue should label it "renewal pending" and not "at risk". A record that has "expired, been deregistered or is otherwise dropped from submission" is removed (www.fda.gov).

What absence means

A record also disappears if FDA inactivates it after a compliance case. For listings, FDA announced that it would inactivate human drug listings that remain uncertified after the renewal period (www.federalregister.gov), and that such listings "may be reactivated with an updated and compliant drug listing submission" (www.federalregister.gov). Auditors add a caution about staleness: registered establishments may have gone out of business without telling FDA (www.gao.gov).

A missing record is therefore ambiguous. It can mean the entity never registered, registered under another name or site, sits outside DECRS scope, let its registration lapse, or was inactivated. The pipeline must keep enough history to tell a record that was present last week from one that never was.

Current Versus Excluded Records

FDA's page links two bulk downloads in zip format: an annual registration status file and a companion labelled "(Excluded)". The page does not define the excluded file, explain why a record lands there, or state the delimiter. This report therefore does not assert its meaning. A related FDA file uses the concept explicitly: in the NDC product file, flag value "U" means "Excluded because the listing certification expired due to lack of listing certification" (www.fda.gov), and FDA publishes a separate NDC excluded-drugs download. A careful pipeline loads the DECRS excluded file as its own table, never merges it silently into the current table, and records the status as "not defined on FDA's page; confirm before relying on it."

Table 1 proposes a status matrix. The states are the pipeline's own classifications, derived from the two files and from the scope exclusions FDA publishes; they are not FDA statuses.

Pipeline state	Evidence in snapshot	What it supports	What it does not support	Default queue action
C1: Current, identifier match	In current file; supplier-provided DUNS or FEI equals file value; normalized address agrees	A current record existed on the snapshot date	Approval or GMP status	Low priority; request registration confirmation and listing evidence
C2: Current, name and address match, no identifier	Exact normalized name and address; no DUNS or FEI supplied	A probable match	Identity of the legal entity beyond the string	Medium; request site DUNS and FEI
C3: Current, address match, name differs	Same normalized address, different name	Possible trade name, parent or ownership change	That the supplier controls the site	High; request legal-entity documents; check the 30-day update rule
C4: Current, fuzzy candidate	Score above a recorded threshold; no exact rule fired	A candidate for a person to review	Any verification	Mandatory human review
X1: In excluded file	In excluded file, not in current	A record exists that FDA's page does not describe as current	Any conclusion about cause	Escalate; ask supplier and FDA
N1: Not found, in-scope claim	No exact or candidate hit in either file	Claim unsupported by this layer on this date	That the supplier is unregistered	Escalate; request registration evidence
N2: Not found, out-of-scope class	Supplier states an excluded class (outsourcing facility, wholesale distributor)	A different register applies	DECRS status	Route to the correct source

Table 1 is a routing device and not a scoring model. One supplier can yield C1 for a manufacturing site and N2 for a distribution arm, and C1, the strongest state this layer can produce, still says nothing about approval. Any state involving a human decision (C3, C4, X1, N1) should leave an auditable note.

Entity and Address Normalization

Matching quality is largely decided before any similarity score. Four deterministic steps are documented in public standards.

- **Unicode normalization.** Canonically equivalent strings can be encoded differently and compare unequal. The Unicode standard defines canonical equivalence as a fundamental equivalency between characters or sequences (unicode.org) and requires that "Normalization Forms remain stable over time" (unicode.org). Python's NFKC form applies compatibility decomposition then canonical composition (docs.python.org); record the form used.
- **Address standardization.** USPS Publication 28 describes "uniform methods for matching addresses with the information in Address Information System" (pe.usps.com) and maps street suffix variants to one official abbreviation (pe.usps.com).

- **Country codes.** ISO 3166 recommends the alpha-2 code for general use (www.iso.org), but "it does not define the names of countries" (www.iso.org), so the pipeline needs its own alias table.
- **Legal-form suffixes.** GLEIF notes that "A legal form only has a meaning according to local legislation" (www.gleif.org). A suffix such as "Pvt Ltd" or "GmbH" can be stripped for a comparison key, but should be stored and not treated as equal across jurisdictions.

Two rules guard against false matches. Keep the raw string beside each normalized key, because the original is the evidence, and never normalize away what distinguishes sites: GS1's location standard says a physical location "must be allocated its own GLN when its access address is" different from others (www.gs1.org). A legal entity's registered address (www.gleif.org) is often a head office and not a plant, which is why FDA asks for a site DUNS and not the headquarters number.

Match Rules: Exact, Fuzzy and Unresolved

Record-linkage research supplies a defensible vocabulary. Deterministic linkage "involves setting decision rules based" on agreement of matching variables (ijpds.org), with rules applied from most to least specific, each looser rule finding more matches "but with increasing risk of false matches" (ijpds.org). Linkage error occurs in two ways, false matches and missed matches (pmc.ncbi.nlm.nih.gov), and there is "an inescapable trade-off" between them (pmc.ncbi.nlm.nih.gov). Where results drive a service decision, "false links can be especially harmful" (pmc.ncbi.nlm.nih.gov), which argues for conservative rules. Exact-only rules lose true matches: in one Census Bureau evaluation, more than 25% of first names and more than 15% of last names among true matches did not agree exactly (www.census.gov). A comparison of algorithms found deterministic rules "typically failed to identify matches in scenarios with low data quality" (pmc.ncbi.nlm.nih.gov).

Uncertain pairs go to people: a methods review describes manual review as a mechanism "often employed to improve probabilistic linkage by subjecting uncertain links" to human assessment (pmc.ncbi.nlm.nih.gov). Measuring error is hard, because linkage errors "do not always occur randomly" (pmc.ncbi.nlm.nih.gov) and representative gold standard data "are rarely available" (pmc.ncbi.nlm.nih.gov). That is why this report declines to publish a match rate.

Edit distance is the common building block. Levenshtein distance counts the insertions, deletions and substitutions that turn one string into another (xlinux.nist.gov), and Jaro-Winkler boosts scores for matching initial characters (xlinux.nist.gov). RapidFuzz returns similarity as a float between 0 and 100 (rapidfuzz.github.io), which makes thresholds explicit and loggable. Source data quality matters too. GAO work described a commercial identifier being requested "to avoid duplications and errors" (www.gao.gov).

Table 2 turns this into a rubric. Each tier uses only evidence reproducible from the snapshot and the supplier's documents; a fuzzy score can raise review priority but cannot lift an item above "candidate".

Tier	Deterministic rule	Evidence required	Allowed wording	Disallowed wording
A: Identifier-exact	Supplier-provided site DUNS or FEI equals the file value, as text with leading zeros preserved; normalized country agrees	Supplier document showing the identifier; snapshot row	"Matched on identifier on snapshot date"	"Verified supplier"

Tier	Deterministic rule	Evidence required	Allowed wording	Disallowed wording
B: Name and address exact	Normalized legal name (suffix stored) and address string equal	Snapshot row; supplier address from contract or certificate	"Exact name and address match; identifier not confirmed"	"Same legal entity"
C: Address exact, name differs	Normalized address equal; name differs after suffix and punctuation rules	Snapshot row; supplier's explanation and ownership documents	"Same site, different name; explanation pending"	"Trade name confirmed"
D: Fuzzy candidate	No exact rule fired; similarity above recorded threshold; country agrees	Score, algorithm, library version, threshold	"Candidate for human review"	"Verified" or "likely match" without a reviewer
U: Unresolved	No candidate, or tied candidates	Search log	"Unresolved; escalate"	"Not registered"

Table 2 is asymmetric by design. Tier A is the only tier that uses an independent key, and it still depends on the supplier's document being genuine, so the queue records where each identifier came from. Tiers B and C use strings, which can coincide by accident (generic names, shared industrial parks). Tier D is never reported as a match without a named reviewer and a reason. A team that publishes match rates should disclose the exact snapshot, the denominator (for example, supplier-site pairs in scope), the rules fired and the items excluded to manual review, and report the unresolved band beside the matched band.

Cross-Checks With NDC, SPL, Inspection and Foreign Registers

The NDC Directory and structured product labeling

The NDC Directory is built from listing files "submitted to FDA in structured product labeling (SPL) electronic listing files by labelers" and is updated daily. Three limits apply. Inclusion "does not indicate that FDA has verified the information provided" (www.fda.gov). APIs and other drugs "not in final marketed form" are outside it (www.fda.gov), so API suppliers cannot be corroborated there, and the directory excludes "drugs manufactured under contract" (www.fda.gov). And FDA "does not submit or alter registration or listing data" (www.fda.gov).

The labeler code identifies a marketing company, not a plant: "A separate labeler code is not required for each establishment manufacturing the drug with that labeler code" (www.fda.gov). A labeler can therefore differ from the maker. For programmatic access, openFDA serves the NDC data (open.fda.gov), limited to 1,000 requests per day per IP address without a key (open.fda.gov), or 120,000 per day with a free key (open.fda.gov); because the directory is updated daily, a pipeline should record the retrieval date alongside any record count it reports. The DailyMed API gives access to current SPL information (dailymed.nlm.nih.gov), though its resources center on NDC codes and not establishments (dailymed.nlm.nih.gov).

Inspection data

FDA's Data Dashboard defines its FEI field as "The FDA Establishment Identifier (FEI) associated with the firm" (datadashboard.fda.gov) and includes an inspection end date (datadashboard.fda.gov). FDA assigns the FEI after initial registration, so the FEI is the natural join key between registration and inspection data, although FDA does not itself call it a join key and that characterization is an inference. Statute directs FDA's registration database to use the facility identifier "to link with other relevant databases" (www.govinfo.gov). The datasets "are updated weekly and only include final actions" (datadashboard.fda.gov). FDA cautions that it "does not issue CGMP certificates at the conclusion of an inspection" (datadashboard.fda.gov). This report interprets no named supplier's inspection history; in a pipeline, inspection data is a dated corroborating field and never an input to automated pass or fail.

Registers outside the United States

Many API sites are outside the United States, so a U.S. record is never the whole picture. Four non-U.S. sources illustrate the range. The EU's EudraGMDP is "the EU database on manufacturing, import and wholesale" authorisations and certificates (www.ema.europa.eu), publicly available since 2011 (www.ema.europa.eu), and one aim is "facilitating the verification of legitimate actors" in the active-substance chain (www.ema.europa.eu). Its content is provided by national competent authorities (eudragmdp.ema.europa.eu), and EMA "accepts no responsibility or liability whatsoever" for it (eudragmdp.ema.europa.eu). The UK's MHRA publishes a monthly register of licensed wholesale sites (www.gov.uk) that reflects data "held by the MHRA on the date of publication" (www.gov.uk), with a site number "allocated to each site" (www.gov.uk). Health Canada requires a drug establishment licence for anyone who will "fabricate, package/label, distribute, import, wholesale or test a drug" (www.canada.ca). WHO's prequalified API list holds sources found acceptable "in principle" (extranet.who.int), yet "Inclusion in the list does not imply any approval by WHO of the products and manufacturing sites" (extranet.who.int), and WHO "cannot guarantee" continued compliance (extranet.who.int). Each has its own scope, identifier and refresh cadence, and none is interchangeable with DECERS.

Discrepancy Triage

Levels of identity (legal entity, site, labeler) and of activity (operation) explain most discrepancies. *Table 3* lays them out.

Level	Typical identifier	Question answered	Public source	Where it diverges
Legal entity / registrant	Legal name; owner or operator; LEI where one exists (www.gleif.org)	Who owns or operates the establishment?	Registration record; supplier's corporate documents	Parent, subsidiary and trade names; one registrant may register many sites
Site / establishment	Site DUNS; FEI; address	Which physical location?	DECERS; inspection data by FEI	One entity, many sites; one address, several firms
Operation	Business operation, e.g. API MANUFACTURE, ANALYSIS	What activity is registered at that site? (www.ecfr.gov)	Registration SPL	A site may be registered for testing only (www.ecfr.gov)

Level	Typical identifier	Question answered	Public source	Where it diverges
Labeler	NDC labeler code	Which firm markets under which code?	NDC Directory	Labeler is not site; private label distributors may hold codes without registering

A "mismatch" is usually a level error. A supplier quoting a holding company names the registrant level while the DECERS row sits at the site level. A contract laboratory can hold a registration limited to analytical operations: FDA's definition of "manufacturer" includes "control laboratories, contract laboratories, contract manufacturers" (www.ecfr.gov), and its operations list separates ANALYSIS from API MANUFACTURE. Auditors have noted that establishments making drugs imported indirectly "may not all register with FDA" (www.gao.gov), so ask where in the chain a claimed API site sits.

Four discrepancy classes follow, each with a default rule.

- **Name differs, address matches.** Ask for ownership or trade-name documents and check whether a name change fell inside the 30-day update window.
- **Address differs, name matches.** Treat as a different site until proven otherwise; ask for the site DUNS.
- **Status differs.** The supplier says registered but the site is absent from the current file or present only in the excluded file. Check the renewal dates first, then escalate without inferring a reason.
- **Operation differs.** Registered operations omit the claimed activity. For an API claim, request registration SPL evidence and quality documents.

Scope routing comes before escalation. FDA's separate outsourcing-facility list was "Updated as of 10/23/2026" and showed 98 entries when retrieved (www.fda.gov); FDA's page does not say that a listing is or is not a verification, so reading an entry on that list as a notice to FDA rather than a certificate from it is this report's own inference, by analogy with registration generally. Wholesale distributors and third-party logistics providers appear in a separate FDA database, and appearing there "does not mean the facility is licensed or approved by FDA" (www.fda.gov); FDA directs buyers to state licensure sites.

Refresh Architecture, Audit Trail and Schema

Architecture

A minimal architecture has six stages, each producing an inspectable artifact.

1. **Acquire.** Download both zip files each business day, since FDA updates the file each business day. Record the retrieval time and the "Data Current through" date. Store timestamps unambiguously: ISO 8601 is the "internationally agreed way to represent dates" (www.iso.org), and RFC 3339 recommends UTC because "true interoperability is best achieved by using Coordinated Universal" Time (www.rfc-editor.org). Treat downloads as untrusted data and extract them to an isolated directory.
2. **Fingerprint.** Compute a SHA-256 digest, available in Python's hashlib (docs.python.org), reading the file in binary mode (docs.python.org). NIST states that such digests "are used to detect whether messages have been changed since the digests were generated" (csrc.nist.gov).

3. **Load raw.** Read all columns as text to keep leading zeros; DuckDB offers an option to "Skip type detection and assume all columns are of type" varchar (duckdb.org). CSV has no single standard: Python's documentation warns of "subtle differences" between producers (docs.python.org), so declare the dialect and encoding explicitly. Store the header row verbatim.
4. **Normalize.** Apply the versioned rules above and write keys beside the raw strings.
5. **Match and queue.** Run the Table 2 tiers for each supplier-site claim and write a queue row for every claim, including clean ones.
6. **Difference.** Compare today's snapshot with the last to detect appearance, disappearance and changes of expiry or address. These differences answer "was this record there last month?".

Do not use the data.gov catalog as a freshness signal: its metadata lists an update frequency of "R/P3M", a three-month pattern (catalog.data.gov), which conflicts with FDA's business-day statement. FDA's page and the search page's data date are better signals.

Audit trail and lineage

Electronic-record rules offer a model even where they do not formally apply to a procurement database. Part 11 calls for "secure, computer-generated, time-stamped audit trails to independently record the date and time" of actions that create, modify or delete records (www.ecfr.gov), and says "Record changes shall not obscure previously recorded information" (www.ecfr.gov). NIST's big data framework says provenance metadata "provides the history of the data so users can correctly interpret the data" (nvlpubs.nist.gov) and recommends linear pathways to authoritative sources "to preserve data lineage" (nvlpubs.nist.gov). In pipeline terms: append-only snapshots, no in-place correction of source rows, and reviewer decisions written as new rows with user, time, reason and snapshot identifier. Quality and compliance staff should decide whether any regulated computerized-system expectation applies.

Downloadable schema

Table 4 gives the pipeline's own field dictionary. Source columns are marked "map at ingest" because FDA does not publish them.

Field	Type	Source	Rule or purpose
snapshot_id	text	Pipeline	Date plus file digest; lineage key
retrieved_at_utc	timestamp	Pipeline	UTC retrieval time (www.rfc-editor.org); used for days-since-refresh
source_current_through	date	Search page	FDA's stated data date
firm_name_raw / key	text	File column, map at ingest; key by pipeline	Original string as evidence; normalized key with suffix kept separately (www.gleif.org)
address_raw / key	text	File columns, map at ingest; key by pipeline	Original and USPS-style key (pe.usps.com)
duns / fei	text	File columns if present, map at ingest	Text, leading zeros kept

Field	Type	Source	Rule or purpose
operations	text	File column if present	Compare with claimed activity
expiry_date	date	File column if present	Days-to-expiry; December 31 pattern
match_tier / rule_id	text	Pipeline	Tier from Table 2; rule that fired
reviewer / decision / reason	text	Human	Append-only

The dictionary separates what FDA provides from what the pipeline derives. The "if present" labels state the documentation gap, and the first build test should fail loudly when a header differs from the previous snapshot.

Case Study: Worked Supplier-Verification Example (Hypothetical Example)

This is a **hypothetical example**: the supplier, plants, identifiers and scores are invented, and only the DECRS data date and retrieval date are real.

A buyer is qualifying an API supplier, "Alderwood Biochem", for a hypothetical intermediate. The quotation names "Alderwood Biochem Holdings Ltd." as seller and lists a plant address. The analyst searched data current through October 26, 2026, retrieved October 27, 2026, a source age of 1 day. *Table 5* shows the queue after the rules are applied.

Claim (hypothetical)	Snapshot evidence (hypothetical)	Tier and state	Days (computed)	Next action	Source date and reference
Plant A, API manufacture	Current row with the same site DUNS as the supplier's certificate; operations include API manufacture; expiry 2026-12-31	Tier A, C1	1 day since refresh; 65 days to expiry, inside the open renewal window	Request confirmation after renewal; no escalation	2026-10-26 data date
Seller: "Alderwood Biochem Holdings Ltd."	Plant A's row carries "Alderwood Chemicals Pvt. Ltd." at the same normalized address	Tier C, C3	1 day	Ask for an ownership chart linking seller to registrant	2026-10-26 data date
Plant B, same group	Candidate scoring 91 with a different street number	Tier D, C4	1 day	Human review; request site DUNS	2026-10-26 data date; (rapidfuzz.gitlab.com/)
Distribution arm, "Alderwood Logistics"	In neither file	N2: out of scope	Not applicable	Route to state licensure and FDA's reporting database	Retrieved 2026-10-27

No row is marked "approved". Row 1 supports only "a registration record with a matching site identifier was present on the snapshot date". Row 2 is the classic entity divergence: three strings for what may be one group, which public data cannot resolve. Row 3 shows that a score of 91 is a reason to look, not an answer. Row 4 shows routing, not failure. The decision to qualify belongs to the quality unit, whose duty of "approving or rejecting drug products manufactured, processed, packed, or held under contract by another company" is written into the cGMP regulation (www.ecfr.gov). ICH Q7 expects a system for evaluating suppliers (www.fda.gov), and FDA's quality agreements guidance places the same approval duty on the owner's quality unit (www.fda.gov).

Data Analysis and Evidence

The quantitative record is useful for scale and as a warning about definitions. FDA's statutory report states that registered drug establishments totalled 13,711 in fiscal year 2025, split into 8,895 domestic and 4,816 foreign. By type, it counts 4,384 finished drug product establishments, 1,951 API establishments and 7,376 "Other" (www.fda.gov). By this report's arithmetic, API establishments are about 14.2% of the total, finished drug product about 32.0%, and foreign about 35.1% (4,816 divided by 13,711). Year-over-year comparisons across annual reports should be made with care because definitions can change.

Several FDA numbers describe different populations. A CDER site catalog counted 5,953 manufacturing sites at the end of fiscal 2025, down 1% from 5,999 in fiscal 2021 (www.fda.gov); that is a different population from the 13,711 registered establishments. The lesson is to name the denominator.

On the global footprint, GAO reported that HHS put the overseas share of API-making establishments at 73 percent as of March 2021 (www.gao.gov).

Registration data also carries duties beyond renewal. The 2020 CARES Act added a duty for registrants to "report annually to the" Secretary on the amount of each listed drug (www.govinfo.gov); teams should confirm current deadlines with FDA.

The pipeline's date arithmetic is also a data product. As of October 27, 2026: days since source refresh = 1; days to a December 31, 2026 expiry = 65; days left in the renewal window = 65. A registration filed on October 1, 2026 would show expiry 2027-12-31, 430 days after October 27, 2026 (65 days to year end plus 365).

Implications and Future Directions

For supplier-quality teams, the main implication is a change of question. "Is the supplier in the FDA database?" becomes "On which dates did which records match which claims, under which rules, and who reviewed the exceptions?" The second question can be answered months later, which is what an audit needs. DECRS stays one evidence layer beside inspection data, quality agreements, audits and supplier documents. Auditors have noted that FDA took steps to improve accuracy by requiring a unique facility identifier at annual registration (www.gao.gov) and by commissioning address validation for selected establishments (www.gao.gov). Identifier-based joining is the intended direction.

Two open items warrant monitoring. The excluded file is undefined on FDA's page, so re-read the page on each release and treat any added definition as a documentation change. And the data.gov metadata and FDA's page disagree on refresh frequency, so freshness should come from the source. Licensure standards for wholesale distributors and third-party logistics providers also vary by state; this report did not confirm the status of any national rule, so teams should check FDA's reporting page and state boards.

The method generalizes to other registers that carry identifiers and dates, such as the foreign registers above. A team should resist collapsing the queue into one supplier score: FDA's regulation says that being assigned a registration number or unique facility identifier does not signal approval (www.ecfr.gov), and a numeric score invites exactly that over-reading.

Frequently Asked Questions (FAQs)

How do I verify a drug manufacturer's FDA registration?

Search DECRS by firm name and note the data date, then use bulk-file snapshots for repeatability. Compare legal entity, site address and registered operations with the supplier's documents and record a tier from Table 2. A record shows registration, not approval.

Can I download the FDA drug establishment registration database?

FDA links a current annual registration status file and a separate excluded file, both zip downloads. The page publishes no column dictionary, so inspect headers at ingest.

How do I look up an FEI number?

FDA points users who need to find or verify an FEI or DUNS to the registration site. Keep FEIs as text, since they have 7 or 10 digits with leading zeros where necessary. The DECRS page does not document whether an FEI is visible in the search output or the zip.

Is there an FDA drug establishment registration data API?

DECRS is offered as a search page and bulk files. For listings, openFDA exposes the NDC Directory (open.fda.gov) and DailyMed offers a web service for current SPL information (dailymed.nlm.nih.gov). Neither substitutes for the registration file.

What does supplier qualification require beyond DECRS?

A quality unit that approves suppliers and contractors, a supplier evaluation system, and contracts that, under ICH Q7, "permit a company to audit its contractor's facilities for compliance with GMP" (www.fda.gov).

Can the check be automated?

Acquisition, normalization, matching and differencing can be automated. The decision cannot: a Tier D candidate is a prompt for review, and no tier is an approval.

Conclusion

DECRS is the right first public check on a drug manufacturing claim and the wrong place to look for a verdict. It shows whether FDA's registration system currently lists a matching establishment and, if the files are snapshotted, whether it did so on a given date. It cannot show that a supplier is approved, inspected, compliant or in scope, and some supplier classes sit outside it by FDA's own description.

The method keeps those limits visible: dated snapshots of both files, deterministic normalization, a tiered rubric that refuses to call a fuzzy score verified, cross-checks with their own dates, and a human review queue with a reason for every decision. Teams that adopt it should publish match rates only with the snapshot, denominator, rules and manual-review exclusions stated, and should consult regulatory counsel on the legal consequences of any specific supplier relationship.

Source links

Links cited in this article, in order of first appearance. Full addresses are provided for readers using extracted text.

1. <https://www.fda.gov/drugs/drug-approvals-and-databases/drug-establishments-current-registration-site-decrs>
2. <https://www.accessdata.fda.gov/scripts/cder/drls/default.cfm>
3. <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-C/part-207>
4. <https://csrc.nist.gov/pubs/fips/180-4/upd1/final>
5. <https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory>
6. <https://datadashboard.fda.gov/oii/api/api-definitions-inspections.htm>
7. <https://www.govinfo.gov/content/pkg/USCODE-2023-title21/html/USCODE-2023-title21-chap9-subchapV-partA-sec360.htm>
8. <https://www.fda.gov/media/192011/download>
9. <https://www.fda.gov/drugs/drug-registration-and-listing-instructions/how-register-drug-establishment>
10. <https://www.fda.gov/media/83219/download>
11. <https://www.dnb.com/en-us/smb/duns/what-is-a-duns-number.html>
12. <https://www.fda.gov/media/131911/download>
13. <https://www.gleif.org/en/about-lei/introducing-the-legal-entity-identifier-lei>
14. <https://www.gsa.gov/about-gsa/newsroom/news-releases/gsa-systems-switch-to-unique-entity-id-sam-on-april-4-2022-03162022>
15. <https://www.govinfo.gov/content/pkg/PLAW-117publ328/html/PLAW-117publ328.htm>
16. <https://www.govinfo.gov/content/pkg/USCODE-2023-title21/html/USCODE-2023-title21-chap9-subchapV-partA-sec352.htm>
17. <https://www.federalregister.gov/documents/2019/08/14/2019-17436/drugs-intended-for-human-use-that-are-improperly-listed-due-to-lack-of-annual-certification-or>
18. <https://www.gao.gov/assets/gao-22-103611.pdf>
19. <https://www.fda.gov/drugs/drug-approvals-and-databases/ndc-product-file-definitions>
20. <https://unicode.org/reports/tr15>
21. <https://docs.python.org/3/library/unicodedata.html>
22. https://pe.usps.com/text/pub28/28c1_001.htm

23. https://pe.usps.com/text/pub28/28apc_002.htm
24. <https://www.iso.org/iso-3166-country-codes.html>
25. <https://www.gleif.org/en/lei-data/code-lists/iso-20275-entity-legal-forms-code-list>
26. <https://www.gs1.org/standards/id-keys/gln/physical-location>
27. <https://www.gleif.org/en/lei-data/access-and-use-lei-data/level-1-data-who-is-who>
28. <https://ijpds.org/index.php/ijpds/article/download/410/2955/1057>
29. <https://pmc.ncbi.nlm.nih.gov/articles/PMC5837697>
30. <https://pmc.ncbi.nlm.nih.gov/articles/PMC6281162>
31. <https://www.census.gov/content/dam/Census/library/working-papers/2005/adrm/rrs2005-02.pdf>
32. <https://pmc.ncbi.nlm.nih.gov/articles/PMC7226047>
33. <https://xlinux.nist.gov/dads/HTML/Levenshtein.html>
34. <https://xlinux.nist.gov/dads/HTML/jaroWinkler.html>
35. <https://rapidfuzz.github.io/RapidFuzz/Usage/fuzz.html>
36. <https://www.gao.gov/assets/gao-10-961.pdf>
37. <https://www.fda.gov/drugs/drug-registration-and-listing-instructions/how-request-labeler-code>
38. <https://open.fda.gov/apis/drug/ndc>
39. <https://open.fda.gov/apis/authentication>
40. <https://dailymed.nlm.nih.gov/dailymed/app-support-web-services.cfm>
41. <https://datadashboard.fda.gov/oii/cd/inspections.htm>
42. <https://www.ema.europa.eu/en/human-regulatory-overview/research-development/compliance-research-development/good-manufacturing-practice/eudragmdp-database>
43. <https://eudragmdp.ema.europa.eu/inspections/logonGeneralPublic.do>
44. <https://www.gov.uk/government/publications/human-and-veterinary-medicines-register-of-licensed-wholesale-distribution-sites>
45. <https://www.canada.ca/en/health-canada/services/drugs-health-products/compliance-enforcement/establishment-licences/drug-establishment-licences.html>
46. <https://extranet.who.int/prequal/medicines/prequalified/active-pharmaceutical-ingredients>
47. <https://extranet.who.int/prequal/medicines/general-information-who-list-prequalified-medicinal-products>
48. <https://www.fda.gov/drugs/human-drug-compounding/registered-outsourcing-facilities>
49. <https://www.fda.gov/drugs/drug-supply-chain-security-act-dscsa/annual-licensure-reporting-wholesale-drug-distributors-and-third-party-logistics-providers>
50. <https://www.iso.org/iso-8601-date-and-time-format.html>
51. <https://www.rfc-editor.org/rfc/rfc3339>
52. <https://docs.python.org/3/library/hashlib.html>
53. <https://duckdb.org/docs/stable/data/csv/overview.html>
54. <https://docs.python.org/3/library/csv.html>
55. <https://catalog.data.gov/dataset/drug-establishments-current-registration-site>
56. <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-A/part-11/subpart-B/section-11.10>
57. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-1r2.pdf>

58. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-10r1.pdf>
59. <https://www.ecfr.gov/current/title-21/chapter-I/subchapter-C/part-211/subpart-B/section-211.22>
60. <https://www.fda.gov/media/71518/download>
61. <https://www.fda.gov/media/86193/download>
62. <https://www.fda.gov/media/188153/download?attachment>
63. <https://www.gao.gov/products/gao-22-103611>
64. <https://www.govinfo.gov/content/pkg/PLAW-116publ136/html/PLAW-116publ136.htm>

THE PUBLISHER

About IntuitionLabs

IntuitionLabs is an AI consulting, custom software development and data engineering firm serving pharmaceutical, biotechnology, medical-device and other life-science organizations. We work with clinical, regulatory, medical-affairs, commercial, quality and IT teams to connect technology decisions with the work people need to accomplish.

AI consulting and adoption

Our AI enablement services cover readiness assessments, use-case selection, governance and policies, team workshops, adoption measurement and ongoing advisory support. We help organizations structure the information layer behind AI: source material, context, permissions and maintained knowledge that make generated answers useful and reviewable. Private LLM inference and hosted AI options support teams evaluating how to operate AI with appropriate control over their data and infrastructure.

Software, data and life-science workflows

IntuitionLabs develops custom software for pharma and biotech, integrates enterprise systems, and builds data engineering and business intelligence solutions. Areas of focus include AI agents, regulatory research, medical writing, medical affairs, CMC information, competitive intelligence and clinical-document workflows. Our eTMF intelligence work includes cross-system reconciliation and inspection-readiness support.

Enterprise platforms and regulated delivery

We provide Veeva services, application support, managed services, integrations and custom applications, alongside enterprise content work involving platforms such as Egnyte. For regulated workflows, our services include GxP enablement, computer-system validation and software development addressing 21 CFR Part 11 requirements. The applicable controls, validation responsibilities and acceptance criteria are defined for each engagement.

Work with IntuitionLabs

Explore [AI enablement](#), [pharma and biotech software development](#), [data engineering and BI](#), and [Veeva services](#). [Contact IntuitionLabs](#) to discuss your workflow, information sources and implementation needs.

IntuitionLabs publishes educational research to help life-science teams make informed technology decisions. Coverage of a product or organization does not imply a client relationship, endorsement or partnership.

Website: <https://intuitionlabs.ai>

This article is educational. Verify consequential medical, legal, financial, regulatory and technical decisions with appropriate professionals and the cited primary sources. Company services are described separately from the article's research findings.