Docket

DOCKET / RESEARCH & PRACTICAL GUIDANCE

CRM vs. Research Database vs. Evidence Store Guide

> Build a more reviewable acquisition pipeline with Docket. Our deal-origination software and managed research help private equity teams screen companies against a mandate and understand the evidence behind each finding.

Published by Docket · 2026-09-22 · crm vs research database vs evidence store · private equity crm · research database · evidence store · deal origination · data provenance · data lineage · field ownership · system of record

Original article: https://docket.capital/articles/crm-research-database-evidence-store



In this article

1. Executive Summary
2. Introduction and Background
3. CRM
4. Research Database
5. Evidence Store
6. Feature Comparison
7. Performance and Benchmarks
8. Data Analysis and Evidence
9. Implications and Future Directions
10. Frequently Asked Questions (FAQs)
11. Conclusion

Executive Summary

A private equity deal origination stack should not force customer relationship management (CRM), target research, and retained evidence into one undifferentiated database. The useful meaning of **single source of truth** is narrower: one authoritative owner for each field, a stable key that connects copies, and an explicit rule for writes, conflicts, freshness, and deletion. CRM should generally own people, relationships, interactions, pipeline stages, and outreach workflow. The research database should own mandate questions, typed answers, unknown and not-found states, reviewer decisions, and screening state. The evidence store should own captured source objects, excerpts, checksums, collection metadata, and version relationships.

That separation follows the nature of the records. Affinity describes private equity pipelines as **multi-stage and non-linear** (affinity.co), while Allvue says its pipeline product tracks contacts, emails, documents, conversations, ideas, and meetings (allvuesystems.com). Those are current vendor claims about workflow scope, not proof that a CRM should become an evidence archive. By contrast, the World Wide Web Consortium's PROV-O standard models entities, activities, and responsible agents, and defines derivation as transforming one entity into another (w3.org). That vocabulary fits source capture, extraction, review, and later revision without requiring Resource Description Framework (RDF).

The architecture therefore needs **field-level authority**, not merely application-level integration. A legal-entity identifier can originate in research, a relationship owner in CRM, an ownership answer in the research system, and the captured page supporting that answer in the evidence store. Copies are permitted, but each copy names its source owner and cannot silently write back. Stable external IDs matter because Salesforce, for example, supports upsert by an external-ID field and returns an error when that identifier matches multiple records (developer.salesforce.com). Merge, split, invalidate, and delete events must preserve aliases or tombstones rather than severing lineage.

Measurement should focus on whether the information system works: **orphan rate**, evidence-link completeness, conflicting active values, and synchronization lag at the 50th and 90th percentiles. Retention is material because a Pew Research Center study of just under **1 million webpages** found **25%** of pages

collected from 2013 through 2023 were inaccessible by October 2023 (pewresearch.org). A live URL alone is not retained evidence. Procurement should test bulk export, identifiers, attachments, timestamps, version history, deletion semantics, and the ability to rebuild relationships outside the vendor.

Introduction and Background

The question **CRM vs research database vs evidence store** is often framed as a product choice. For an origination leader or data architect, it is more usefully framed as a record-ownership decision. Each system performs a different job, and the handoff between them determines whether a team can reconstruct why a target advanced, which evidence supported an answer, and what changed after review.

Private equity CRM pages emphasize connected relationship and pipeline work. Affinity says relationship tracking should persist across funds (affinity.co), and Allvue describes relationship insights across accounts, contacts, opportunities, and products (allvuesystems.com). These claims describe a natural home for relationship and process state. They do not establish that a mutable contact or opportunity record is the right home for captured web pages, prior conclusions, or evidence-review history.

Docket, the publishing brand, is a direct provider of sourced target research rather than a general-purpose CRM. Its platform says each researched answer includes a source, supporting excerpt, and collection date, and its existing evidence guide argues that a newer observation should become a new record while the current view is computed. This report does not repeat that guide's discussion of source quality. It operationalizes the next architectural question: which system owns which field, what may be copied, and what must remain linked.

The recommended unit of governance is therefore the **field plus its lifecycle**. For each object, define an authoritative system, stable key, permitted writer, reviewer, downstream copies, provenance pointer, freshness rule, conflict rule, and retention rule. The result can still appear unified to users through views and integrations. Unity of presentation does not require unity of storage.

CRM

Capabilities

A CRM is strongest where the record changes through human interaction. Its core objects are organizations, people, relationships, communications, tasks, opportunities, and pipeline events. HubSpot, for example, defines contacts, companies, deals, and tickets as standard CRM objects (developers.hubspot.com). Microsoft says Dynamics 365 dashboards can show opportunity counts at each pipeline stage (learn.microsoft.com). These examples support a practical boundary: CRM owns the relationship and operating process around a target, not every fact known about it.

In a private equity deal origination CRM, authoritative fields normally include:

- **Contact identity:** person name, current role, verified communication coordinates, and CRM record ID.
- **Relationship state:** relationship owner, introducer, strength classification, restrictions, and last substantive interaction.
- **Activity:** emails, meetings, calls, notes, tasks, and their timestamps.

- **Pipeline:** stage, stage-change time, responsible team member, next action, and process status.
- **Outreach control:** approved channel, suppression state, approval status, and message version reference.

Allvue says its Outlook integration associates emails with opportunities and synchronizes contacts (allvuesystems.com). That is a sensible CRM responsibility because the activity belongs to a relationship workflow. A research answer such as ownership type may be displayed in the CRM, but the display should be a read-only projection with an answer ID, review state, observed date, and evidence link.

Adoption and operating fit

CRM adoption is not the same as architectural authority. Users may prefer one screen, and a CRM can host custom fields, but convenience does not decide which system may overwrite a reviewed fact. Affinity describes automatic capture of email and calendar data that creates or updates contact and company records (affinity.co). This supports activity capture, yet enrichment and evidence review still require separate authority rules.

The CRM should carry foreign keys such as `target_id`, `legal_entity_id`, and `current_answer_id`. It may cache summary fields for search and segmentation. Every cache needs a source system, source version, last synchronized timestamp, and refresh status. Salesforce's use of external IDs for upserts illustrates the stable-key pattern, while Microsoft Dataverse documents alternate keys based on business columns rather than only platform-generated identifiers (learn.microsoft.com).

Strengths and limitations

The CRM's strengths are workflow visibility, relationship context, activity capture, assignment, and pipeline reporting. Its main architectural limitation is mutability. A sales-style record usually represents the latest operational view, whereas research needs multiple observations, explicit unknowns, and reviewer decisions. Salesforce opportunity history records change details and the user responsible (help.salesforce.com), but field history is not a substitute for storing the captured source object and the derivation from source to answer.

Merge behavior also requires caution. HubSpot says merged records cannot be unmerged (knowledge.hubspot.com). The integration layer should therefore retain aliases and issue an explicit merge event. A mistaken merge may require a later split event that creates new canonical IDs and reassigns relationships without rewriting the historical evidence.

Research Database

Capabilities

The research database owns the questions the mandate asks and the structured conclusions produced from evidence. Its basic unit is not a company row with many loosely defined columns. It is a typed answer connected to a target, question version, observation period, research run, reviewer decision, and one or more evidence references.

Required states should include:

- **Answered:** a value exists and meets the question's evidence standard.
- **Unknown:** the available material does not permit a supported answer.
- **Not found:** named sources were checked on a stated date without locating the requested evidence.
- **Not applicable:** the question does not apply under the current mandate definition.
- **Pending:** collection or review has not finished.
- **Conflicted:** active observations disagree and no resolution has been accepted.

JSON Schema makes an important distinction for implementation: omitting a property is valid by default (json-schema.org), and an explicit null is not equivalent to an absent property (json-schema.org). A research model must therefore define its states rather than treating blank, null, and not-found as interchangeable.

Adoption and operating fit

A private equity research database fits work that applies the same mandate to many targets. It supports consistent definitions, coverage reporting, bulk refresh, independent review, and comparisons among companies. Commercial research platforms illustrate how typed data and relationships can be delivered downstream. Crunchbase documents entity lookup that returns data fields and relationships for a specified entity (data.crunchbase.com), while Dealroom documents API access to companies, funding rounds, jobs, accelerators, and investors (knowledge.dealroom.co). These are first-party descriptions of scope, not claims that either database satisfies a firm's evidence standard.

The research system should assign immutable IDs to questions, answer observations, review decisions, and runs. A mutable mandate label such as `founder_owned` is not enough because its definition may change. Store `question_version_id` with every answer. When the mandate changes, the system can queue a refresh rather than silently reinterpreting an old answer under a new rule.

Strengths and limitations

The research database is strongest at typed answers, repeatable coverage, explicit missingness, and reviewer-controlled current state. Docket says unresolved questions remain visible after evidence review, and its platform presents coverage across answered, not-found, and pending questions. That is the right pattern for research completeness.

Its limitation is that a structured conclusion is not the source itself. Even a reviewed ownership answer can outlive or lose its linked page. The research database may keep excerpts for query performance, but it should point to immutable evidence objects. It also should not become the authoritative owner of relationship notes or outreach approvals merely because those fields are convenient to show beside the research.

Evidence Store

Capabilities

An evidence store retains the material that was observed and the context required to authenticate and revisit it. Its core objects include source capture, content version, excerpt, checksum, collection activity, collector identity, and relationships such as derived-from, revised-from, and invalidated-by.

The W3C provenance model supplies a useful subset. An entity is a physical, digital, conceptual, or other thing with fixed aspects (w3.org). An activity occurs over time and acts on or with entities (w3.org). An agent bears responsibility for an activity (w3.org). A practical implementation can call these **artifact**, **run**, and **actor** without using RDF.

Evidence objects should be immutable. Google Cloud Storage documents that uploaded objects do not change and that the combination of object name and generation uniquely identifies a version (docs.cloud.google.com). It also supplies a CRC32C hash for every object (docs.cloud.google.com). These product details illustrate general design properties, not a requirement to use that vendor.

Adoption and operating fit

The evidence store fits audit, re-review, refresh, and reconstruction. A source capture should include the requested URL, resolved URL, retrieval time, content type, extraction method, content hash, access classification, and retention class. The excerpt becomes its own addressable object with offsets into the captured content. The accepted answer references the excerpt and capture, never only the live URL.

This architecture matters because live content disappears. Pew found that **38%** of pages in its 2013 snapshot were no longer accessible in 2023 (pewresearch.org). Even among government pages examined, **21%** contained at least one broken link (pewresearch.org). Retention does not prove a conclusion, but it preserves the material the reviewer actually saw.

Strengths and limitations

The evidence store's strength is point-in-time fidelity. It supports integrity checks, historical reconstruction, access control, and later reprocessing. Its limitation is that it should not decide the current business meaning of evidence. The accepted ownership classification belongs to the research system; the relationship owner's next action belongs to the CRM.

Immutability also does not mean indefinite retention. NIST says audit-record retention should align with an organization's records-retention policy (nvlpubs.nist.gov), while the UK Information Commissioner's Office says a retention schedule should list the types of records held (ico.org.uk). Security, privacy, contractual, and jurisdictional requirements remain firm-specific. Immutability within a retention period must coexist with authorized deletion or anonymization after that period.

Feature Comparison

Table 1 compares the three architectural roles and places Docket according to its actual model. Docket is a direct provider of target research and retained evidence functions, not a substitute for a firm's CRM.

Option	Authoritative scope	Primary writers	Best query	Important limitation
CRM	Contacts, relationships, interactions, pipeline events, tasks, outreach approvals	Origination team and connected communication systems	Who knows this target, what happened, and what happens next?	Mutable workflow records do not by themselves preserve source context.

Option	Authoritative scope	Primary writers	Best query	Important limitation
Research database	Mandate questions, typed answers, unknowns, review decisions, screening state	Researchers, enrichment processes, and reviewers under explicit rules	Which targets meet the current mandate, with what confidence and freshness?	Structured answers still require durable links to the observed material.
Evidence store	Source captures, excerpts, hashes, collection metadata, versions, derivation links	Controlled capture and preservation services	What material supported this answer at the time?	It preserves observations but should not own relationship or pipeline judgment.
Docket	Sourced target research with answer, excerpt, source, collection date, separate review, and export	Triage, Scout, Audit, and authorized users within its research workflow	What supports each acquisition-target answer, and which questions remain open?	It complements rather than replaces the firm's relationship CRM. Docket says results can be exported for use outside the platform.

The comparison shows why “CRM for acquisition target research” is an integration pattern, not a reason to collapse the categories. A user can work from a CRM screen while the authoritative research answer and immutable source artifact remain elsewhere. Conversely, an evidence-backed research workspace can show pipeline context without becoming the authoritative owner of emails and relationship notes.

Field ownership and synchronization design

Table 2 is a RACI-like field-ownership matrix. “Writer” indicates the role allowed to change the authoritative value. Reviewers may accept or reject a proposed value without directly editing the source artifact.

Field or object	Authoritative system	Stable key	Writer	Reviewer	Downstream copies	Provenance link	Freshness rule	Conflict rule	Retention
Legal entity	Research database	`legal_entity_id`	Identity resolver	Research reviewer	CRM account, reports	Identity-source capture IDs	Refresh on material identity signal	Quarantine ambiguous matches	Keep identity history while target record exists
Brand or trading name	Research database	`brand_id` plus entity link	Research process	Research reviewer	CRM display name	Source and observation IDs	Recheck on website or filing change	Allow dated aliases	Follow research-record schedule
Contact	CRM	`contact_id` plus external aliases	CRM users and approved sync	Relationship owner	Research workspace	Source reference for enriched attributes	Verify role by field policy	Human review before merging people	Relationship policy and applicable privacy rules

Field or object	Authoritative system	Stable key	Writer	Reviewer	Downstream copies	Provenance link	Freshness rule	Conflict rule	Retention
Relationship	CRM	`relationship_id`	Relationship owner	Origination lead	Research view	Interaction IDs	Event-driven	Manual resolution of competing owners	Firm policy
Interaction	CRM	`interaction_id`	Connected channel or user	Relationship owner	Research timeline view	Native activity reference	Event-driven	Append correction, do not replace history	Communications policy
Target	Research database	`target_id`	Research operations	Mandate owner	CRM opportunity/account	Entity and mandate IDs	Recompute on mandate change	One canonical target, aliases retained	Mandate plus archival policy
Mandate answer	Research database	`answer_observation_id`	Research process	Independent reviewer	CRM summary and analytics	Excerpt and capture IDs	Per-question expiry	Preserve competing observations, select reviewed current view	Preserve superseded versions per policy
Source capture	Evidence store	`capture_id` plus content hash	Capture service	Evidence reviewer	Metadata only	Original URL and activity ID	Immutable after capture	New capture for new content	Evidence retention class
Excerpt	Evidence store	`excerpt_id`	Extraction process	Evidence reviewer	Research answer view	Capture ID and offsets	Immutable; supersede by new excerpt	Reject if offsets or hash fail	With parent capture
Review decision	Research database	`review_id`	Reviewer	Review lead for exceptions	CRM acceptance flag	Answer and evidence IDs	New decision for changed evidence	Latest valid decision governs current view	Full decision history
Pipeline event	CRM	`pipeline_event_id`	CRM workflow	Deal team	Research context view	Prior and new stage IDs	Event-driven	Reject stale sequence; preserve event	CRM audit policy
Outreach approval	CRM	`approval_id` plus content version	Authorized approver	Origination lead	Outreach service	Target, rationale, and message version IDs	Invalidate when content changes	No send if approval version differs	Communications policy

No downstream copy in the matrix gains authority merely by being easier to edit. NIST recommends analyzing and correlating audit records across repositories (nvlpubs.nist.gov). The same principle applies here: correlate histories, but preserve which repository authored each event.

Write-back rules

Write-back should be asymmetric and explicit:

- **Analyst judgment may change research state:** a reviewer can accept an ownership answer or move a target from pending to screened, creating a new decision record.
- **CRM workflow may consume, not rewrite, research:** a reviewed answer may update a CRM cache, but a CRM user edit cannot overwrite the accepted research observation.
- **Enrichment proposes rather than replaces:** a new value creates a candidate observation. It does not silently replace a reviewed fact.
- **Evidence is append-only within retention:** corrected extraction creates a new excerpt and a supersession link. The captured object remains unchanged.
- **Pipeline writes remain in CRM:** a changed stage creates a pipeline event and may trigger research work, but it does not recast the evidence.
- **Approval binds to a version:** changing an approved outreach message invalidates the approval tied to the earlier version.

HTTP's If-Match precondition exists to prevent accidental overwrites by parallel actors (rfc-editor.org). An equivalent optimistic-concurrency rule should protect each authoritative update: write only if the version last read is still current.

Performance and Benchmarks

There is no public benchmark that establishes one universal latency, completeness, or accuracy target for this three-system architecture. Performance should be tested against the firm's workflow and service objectives. The useful benchmark is whether every accepted answer remains traceable, every authoritative change reaches required consumers, and failures become visible before users act on stale state.

Event contract

Table 3 defines the minimum events required to synchronize without granting every system write authority.

Event	Required payload	Consumer behavior	Idempotency and conflict rule
Create	Event ID, object type, canonical ID, version, occurred time, owner system, minimal fields	Create projection or fetch authoritative record	Ignore an already applied event ID; reject duplicate canonical ID with different identity
Update	Object ID, prior version, new version, changed fields, actor, occurred time	Apply only to downstream cache; retain source version	Compare expected prior version; queue gaps for replay
Invalidate	Object ID, invalidating activity, reason code, replacement ID if any	Stop presenting value as current; keep history	Never translate invalidation into hard deletion

Event	Required payload	Consumer behavior	Idempotency and conflict rule
Merge	Surviving ID, alias IDs, relationship reassignment map, actor, occurred time	Redirect aliases and reattach projections	Keep alias lookup permanently or for documented retention period
Split	Former canonical ID, new IDs, allocation map, reviewer decision	Create new identities and reassign future projections	Do not rewrite evidence captured under the former understanding
Delete	Object ID, deletion class, effective time, retention or legal-policy reference	Remove or anonymize authorized copies; retain permitted tombstone	A deletion replay must remain idempotent and auditable

CloudEvents requires the combination of event source and ID to be unique for each distinct event (github.com). HTTP also defines PUT, DELETE, and safe request methods as idempotent (rfc-editor.org). Together these patterns support safe replay. Ordering still cannot be assumed: Azure Event Grid warns that subscribers might receive events out of order (learn.microsoft.com). Consumers should compare object versions or sequence numbers before applying an update.

Worked ownership refresh (Hypothetical Example)

Consider a fictional target, **Northbridge Controls**, researched for a fictional industrial mandate. The example illustrates state transitions and makes no claim about an actual company.

1. **Collection:** Scout captures a company page as `capture_884`, records its hash and collection time, and creates `excerpt_191` pointing to the relevant span.
2. **Research:** the research database creates `answer_620`, question version `ownership_v3`, value `founder_owned`, state `pending_review`, and a derivation link to `excerpt_191`.
3. **Review:** a reviewer accepts the answer in `review_074`. The computed current answer becomes `answer_620` without changing the source capture.
4. **CRM projection:** an update event writes `ownership_summary=founder_owned`, `answer_id=answer_620`, `review_state=accepted`, and `observed_at` into read-only CRM fields.
5. **Later refresh:** a new source supports a different ownership state. The system creates a new capture, excerpt, and answer observation. It does not edit `answer_620`.
6. **Conflict handling:** if the two sources cover overlapping dates and disagree, the new answer remains conflicted until review. The CRM continues showing the last accepted answer with a stale or conflict flag.
7. **Resolution:** the reviewer accepts the newer observation. A new event updates the CRM projection and retains the earlier answer for historical reconstruction.

PROV-O defines revision as a resulting entity that is a revised version of an original (w3.org). The example uses the same idea while keeping observations immutable and computing which one is current.

Data Analysis and Evidence

The architecture should be evaluated with reconciliation measures, not anecdotal confidence or claimed deal outcomes. NIST's quality-of-data framework expresses completeness as the proportion of stored data with complete information against a potential 100% (nist.gov). Origination teams can adapt that structure to four

operational measures:

- **Orphan rate:** authoritative records whose required parent or stable-key mapping is missing, divided by all records tested. Report separately for answers without targets, excerpts without captures, and CRM projections without authoritative source IDs.
- **Evidence-link completeness:** accepted answers with a resolvable excerpt and retained capture, divided by all accepted answers that require quotable evidence. Do not count a live URL alone as complete.
- **Conflicting active values:** object-field pairs with more than one active authoritative value and no recorded conflict state. The target should be zero because legitimate disagreement belongs in explicit competing observations.
- **Synchronization lag p50 and p90:** elapsed time from authoritative event occurrence to verified downstream application at the median and 90th percentile. Publish both because a median can hide a slow tail.

These measures should be sliced by object type, source system, consumer, and event type. A single global percentage can conceal that contact synchronization is healthy while evidence links are incomplete. Counts should accompany percentages so a small denominator remains visible.

Delivery documentation provides useful test conditions, not universal targets. Microsoft Graph treats delivery as successful after a 2xx response within **3 seconds** and retries unsuccessful notification delivery for up to **4 hours** (learn.microsoft.com). GitHub asks webhook endpoints to respond within **10 seconds** and permits manual redelivery of events from the prior **3 days** (docs.github.com). These differences show why the integration layer needs a durable event log and its own replay policy rather than assuming the provider will retain events forever.

Before launch, run a controlled migration and reconciliation test:

- **Seed:** load a known set of entities, contacts, answers, artifacts, merges, and deletions.
- **Count:** compare row and object counts by type at source and destination.
- **Trace:** sample records from destination back to stable source IDs and retained evidence.
- **Replay:** resend duplicate and out-of-order events and confirm idempotent results.
- **Conflict:** attempt a write based on an outdated version and confirm it is rejected or queued.
- **Delete:** execute each deletion class and verify downstream copies, aliases, and tombstones.
- **Refresh:** introduce newer evidence and confirm that the reviewed fact is not silently overwritten.

Implications and Future Directions

The most important implication is organizational: **system ownership must be assigned at field level**. Saying “the CRM is the system of record” is incomplete when the same target has relationship data, mandate judgments, and captured evidence with different writers and retention needs. Architecture documentation should name the owner for each object and the direction of every copy.

Access control follows the same boundary. Relationship notes may be visible to a deal team but not to every research contractor. Evidence artifacts may contain licensed or personal information with narrower access and retention. NIST's least-privilege guidance calls for allowing only authorized access needed for users' tasks (nvlpubs.nist.gov). Apply permissions to artifact, answer, interaction, and export classes rather than copying all data into the least restrictive tool.

Vendor exit should be designed before procurement. A portability review should verify:

- **Bulk export:** all objects, not only a filtered report. Dealroom's documentation, for example, describes CSV export of search results (knowledge.dealroom.co).
- **Identifiers:** canonical IDs, external IDs, merge aliases, and relationship keys. HubSpot says full-property exports put Record ID in the first column (knowledge.hubspot.com).
- **Attachments:** original files, metadata, versions, and links between files and records. Jira Cloud documentation says its native CSV export does not bulk-download physical attachments (support.atlassian.com).
- **Timestamps:** created, observed, modified, reviewed, synchronized, and deleted times with time zones.
- **Change history:** prior values, actors, reasons, and sequence or version markers. HubSpot says merged property histories are combined in the final record (knowledge.hubspot.com).
- **Deletion semantics:** hard deletion, anonymization, invalidation, tombstones, recovery windows, and propagation. The W3C Activity Streams primer explains that a tombstone lets a reference remain in collections after deletion (w3.org).
- **Rebuild test:** reconstruct relationships and current views in an empty environment using only exported material and documented schemas.

An export is not necessarily a restorable backup. Notion explicitly warns that exported workspace content cannot instantly recreate the workspace (notion.com). Procurement should therefore test restoration, not merely confirm that an export button exists.

Frequently Asked Questions (FAQs)

What should a private equity deal origination CRM own?

A **private equity deal origination CRM** should own contacts, relationship responsibility, interactions, tasks, pipeline state, and outreach approvals. It may display reviewed target facts, but those copies should reference authoritative research answer IDs and remain read-only.

What belongs in a private equity research database?

A **private equity research database** should own mandate definitions and versions, typed answer observations, unknown and not-found states, review decisions, target screening state, freshness policy, and links to evidence. It should not overwrite immutable source captures.

What is deal origination data management?

Deal origination data management is the governance of objects, field authority, identifiers, synchronization, review, access, retention, and export across the origination lifecycle. It is broader than choosing a CRM and narrower than a firm-wide data-governance program.

How should a CRM for acquisition target research connect to evidence?

A **CRM for acquisition target research** should store a projection containing the current reviewed value, answer ID, observation date, review state, and evidence link. Selecting the link should open the retained excerpt and source capture in the research or evidence system.

What makes evidence-backed deal sourcing different?

Evidence-backed deal sourcing preserves the connection from a screening judgment to the observed material, collection activity, and reviewer decision. It keeps conflicting or superseded observations visible instead of treating the latest copied value as self-proving.

Where to store deal origination data?

For teams asking **where to store deal origination data**, the answer depends on the field: relationships and pipeline in CRM, mandate answers and review state in the research database, and captured artifacts and excerpts in the evidence store. A shared identifier layer connects them.

What is a private equity evidence store?

A **private equity evidence store** is a controlled repository for point-in-time source captures, excerpts, hashes, timestamps, access classification, and provenance relationships supporting target research. It is not a replacement for CRM workflow or research judgment.

Deal sourcing CRM vs database: which is the system of record?

For **deal sourcing CRM vs database**, neither should own every field. CRM is authoritative for relationship and pipeline records. The research database is authoritative for structured mandate answers and decisions. The evidence store is authoritative for retained source objects. The cross-system contract, not a universal database, creates the reliable record.

Conclusion

The correct choice is not CRM or research database or evidence store. A defensible origination architecture uses all three roles and assigns each field to exactly one authoritative owner. CRM manages people, relationships, activity, pipeline, and approvals. The research database manages questions, answers, unknowns, review decisions, and current screening state. The evidence store preserves the source objects and provenance needed to reconstruct those answers.

Copies are useful when their direction is documented. Every copy should carry a stable source ID, version, observed or synchronized time, and owner system. Write-back is allowed only under an explicit contract. Enrichment creates a candidate observation, review creates a decision, and a later source creates a new

version. None should silently overwrite a reviewed fact.

Implementation quality should be judged by reconciliation evidence: orphan rate, evidence-link completeness, conflicting active values, and synchronization lag at p50 and p90. Migration should test merges, splits, replay, deletion, attachments, history, and restoration from export. Those controls do not promise better deal outcomes. They establish something more basic and testable: whether an origination team can tell which system owns a fact, what supported it, who changed it, and whether every downstream copy is still current.

Source links

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

1. https://www.affinity.co/guides/how-to-choose-the-best-private-equity-crm-software
2. https://www.allvuesystems.com/solutions/pipeline-management
3. https://www.w3.org/TR/prov-o
4. https://developer.salesforce.com/docs/platform/api-rest/guide/dome-upsert.html
5. https://www.pewresearch.org/data-labs/2024/05/17/when-online-content-disappears?trk=article-ssr-frontend-pulse_little-text-block
6. https://developers.hubspot.com/docs/api-reference/latest/crm/objects/schemas/guide
7. https://learn.microsoft.com/en-us/dynamics365/sales/dashboards
8. https://learn.microsoft.com/en-us/power-apps/developer/data-platform/define-alternate-keys-entity
9. https://help.salesforce.com/s/articleView?id=sales.opp_history.htm&language=en_US&type=5
10. https://knowledge.hubspot.com/records/merge-records?app=1
11. https://json-schema.org/understanding-json-schema/reference/object
12. https://data.crunchbase.com/v4-legacy/docs/using-entity-lookup-apis
13. https://knowledge.dealroom.co/knowledge/dealroom-api-1
14. https://docs.cloud.google.com/storage/docs/objects
15. https://docs.cloud.google.com/storage/docs/metadata
16. https://nvlpubs.nist.gov/nistpubs/SpecialPublications/800-171r3/NIST.SP.800-171r3.html
17. https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/data-protection-principles/a-guide-to-the-data-protection-principles/storage-limitation
18. https://www.rfc-editor.org/rfc/rfc9110.html
19. https://github.com/cloudevents/spec/blob/main/cloudevents/spec.md?plain=1
20. https://learn.microsoft.com/en-us/azure/event-grid/delivery-and-retry?trk=article-ssr-frontend-pulse_little-text-block
21. https://www.nist.gov/itl/health-it-testing-infrastructure/testing-tools/qdar-quality-data-rest
22. https://learn.microsoft.com/en-us/graph/change-notifications-delivery-webhooks
23. https://docs.github.com/en/webhooks/using-webhooks/best-practices-for-using-webhooks
24. https://knowledge.dealroom.co/knowledge/how-to-create-and-save-searches
25. https://knowledge.hubspot.com/import-and-export/export-records
26. https://support.atlassian.com/jira/kb/how-to-export-issues-from-jira-cloud-in-csv-format

27. https://www.w3.org/wiki/Activity_Streams/Primer/Tombstone_type
28. https://www.notion.com/help/workspace-settings

THE PUBLISHER

About Docket

Docket provides deal-origination research software and managed research for private equity firms. We help investment teams investigate acquisition targets using structured screening criteria, retained sources and reviewable company evidence. Teams can work through a self-serve platform or use managed research, depending on how they want research delivered.

Research against a defined mandate

A useful target list needs more than company names. Docket focuses on the questions that determine whether a company fits an investment mandate, including the evidence needed to support or qualify each answer. Our research approach makes the connection between screening criteria, source material and conclusions visible to the team reviewing the work.

Triage, Scout and Audit

Docket's three named research agents perform complementary tasks. Triage resolves company identity and screens fit. Scout collects sourced answers against the mandate. Audit checks retained evidence, addresses contradictions and leaves unsupported answers visibly unresolved. This structure helps reviewers distinguish established findings from missing information and questions requiring further investigation.

Evidence that supports investment-team judgment

Our research library covers market mapping, screening criteria, private-company data, succession and ownership, source evaluation and evidence standards. These resources explain the methods and limitations behind origination research. Findings support a team's judgment; they do not establish that a company is for sale or guarantee a transaction or investment outcome.

Work with Docket

Visit Docket to explore the platform, managed research and the current contact path. Read about screening criteria, evidence standards, AI research agents and market mapping.

Public examples are illustrative unless explicitly identified otherwise. Research preparation and authorized outreach are separate activities; confidential target lists and customer outcomes should never be inferred from an educational example.

Website: https://docket.capital

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.