

Data Migration — Scope & Approach

Platform: HubSpot → Salesforce | **Type:** Migration Planning | **Prepared by:** RevPal

Migration Principles

1. **Selective, not bulk.** Only migrate data that will be actively used in Salesforce. Do not replicate HubSpot's data quality problems.
2. **Clean before import.** All data passes through a staging CSV with validation before entering SF.
3. **Test first.** Full migration runs against the Full Copy Sandbox before production.
4. **Preserve attribution.** Maintain Lead Source, partner attribution, and owner assignments where they exist.

Data Volume Context

HubSpot currently holds ~23,000 contacts — a relatively small dataset that simplifies migration planning. Volume estimates below reflect the subset that meets migration criteria.

Migration Method

RevPal will handle all data extraction using programmatic tools (API exports, bulk queries). The client team does not need to export flat files from HubSpot manually.

Migration Progress (as of 2026-03-14)

Completed Steps

Step	Date	Details
HubSpot Contact Export	2026-03-14	23,948 contacts exported via HubSpot CRM export with all properties
Company Association Export	2026-03-14	v4 Batch Associations API → 25,088 rows, 20,712 contacts (86.5%) linked to 17,898 unique companies
Deal Association Export	2026-03-14	v4 Batch Associations API → 1,780 rows, 1,661 contacts (6.9%) linked to 1,605 unique deals
Company Web Enrichment	2026-03-14	14,381 companies enriched with scraped addresses, titles, descriptions, LinkedIn, industry classification
SF Account Creation	2026-03-14	13,330 Accounts created from enriched HubSpot companies + 86 newly-discovered companies (Bulk Job: 750WD00000SKIfNYAT)
SF Account ID Cross-Reference	2026-03-14	13,320 SF Accounts matched to HubSpot Company IDs via <code>HubSpot_Company_ID__c</code> field
Contact-to-Account Enrichment	2026-03-14	18,619/23,948 contacts (77.7%) matched to SF Account IDs; 2,093 have HS Company ID but no SF match; 3,236 have no company association

Account.Type Backfill	2026-03-14	13,324 Prospect, 3 Customer, 2 Partner — all typed by Record Type
Qualification Status Field	2026-03-14	<code>Qualification_Status_c</code> picklist deployed on Account for data quality triage during import
403-Blocked Domain Scraping	2026-03-14	Puppeteer headless browser scraped 1,276 domains that returned 403 Forbidden — 775 addresses extracted (61% success rate), 3-pass address cleaning applied
Company Name Enrichment	2026-03-14	1,043 company names derived: 841 from existing page titles + 193 from newly-scraped titles + 9 from og:site_name/JSON-LD
Data Quality Disqualification	2026-03-14	1,087 accounts marked "Disqualified - Out of Business" (unreachable sites + 404s), 26 marked "Disqualified - Not in ICP" (free email domains), 117 marked "Disqualified - Not in SAM" (non-US companies across 20 countries)
Address Upload to SF	2026-03-14	4,052 billing addresses uploaded (4 failures: scraped content exceeded 255-char BillingStreet limit)
Company Name Upload to SF	2026-03-14	861 company name updates uploaded (filtered from 975 to remove scraped page titles masquerading as names)
Qualification Status Upload to SF	2026-03-14	1,113 qualification statuses uploaded + 117 non-US "Not in SAM" updates

Sample Data Cleanup	2026-03-14	73 Salesforce sample records deleted (3 Accounts, 2 Contacts, 6 Leads, 50 Opportunities, 6 Cases, 6 Tasks)
Account Dedup Fields	2026-03-14	Duplicate_Survivor__c (Checkbox) and Duplicate_IDs__c (LongTextArea 131072) deployed to staging with FLS on Client_RevOps_Admin and Client_Account_Fields
Account Dedup Analysis	2026-03-14	13,415 accounts analyzed via domain + name matching with Union-Find clustering. 18 duplicate clusters found (36 accounts), all REVIEW — different HubSpot IDs on same domain. 0 regional office hierarchies detected. Pending human review.

Actual Volumes (vs. Original Estimates)

Data Set	Original Estimate	Actual Volume	Notes
Contacts in HubSpot	~23,000	23,948	Full export complete
Companies (HubSpot)	Not estimated	17,898 unique	Via contact-company associations
Deals (HubSpot)	Not estimated	1,605 unique	Via contact-deal associations
Accounts in SF (staging)	~200-400	14,467	13,330 enriched + 86 new + pre-existing; 1,230 disqualified (1,087 OOB + 26 Not in ICP + 117 Not in SAM)
Contacts with SF Account match	~500-1,000	18,619	77.7% of total — far exceeds Tier 1 estimate
Contacts without Account	Excluded	3,236	No HubSpot company association
Contacts with unmatched company	N/A	2,093	HubSpot company not in SF — under review

Key Discovery: No HubSpot-Salesforce Sync

The Acme SaaS HubSpot portal ([PORTAL-ID]) has **no native Salesforce sync configured**. The `salesforceid` property does not exist on Company objects. Cross-platform matching relies entirely on the `HubSpot_Company_ID__c` custom field uploaded to SF Account records during the enrichment phase.

Remaining Steps

Step	Status	Dependency
Contact qualification review	Not started	Qualification criteria from client stakeholder
Contact import to SF	Not started	Qualification review + owner mapping
Deal/Opportunity import	Not started	Contact import (for Contact Roles)
Owner mapping (HubSpot → SF)	Not started	Active SF user list
Demand Funnel Stage backfill	Not started	Contact import
Post-migration validation	Not started	All imports complete

What Migrates

Tier 1: Required at Go-Live

Data Set	Source	Volume Estimate	SF Target Object	Notes
Active customer accounts	HubSpot Companies	~200-300 (actual: 3)	Account (Customer RT)	Only companies with active deals or customer status
Prospect accounts with active pipeline	HubSpot Companies	~50-100 (actual: 14,464)	Account (Prospect RT)	Broader import — all HubSpot companies with enrichment data
Contacts on active accounts	HubSpot Contacts	~500-1,000 (actual: 18,619 matchable)	Contact	18,619 contacts matched to SF Accounts; 23,948 total exported
Open, actively worked pipeline	HubSpot Deals	~35-50	Opportunity	Deals past Qualify with activity in last 90 days — 1,605 unique deals identified
Closed-Won deals (last 12 months)	HubSpot Deals	~50-80	Opportunity	Amount > 0, for historical reporting

Tier 2: Migrated Post-Go-Live (Week 2-4 after cutover)

Data Set	Source	Volume Estimate	SF Target Object	Notes
Closed-Lost deals (last 12 months)	HubSpot Deals	~150-200	Opportunity	For win/loss analysis. Lower priority than Tier 1.
Historical contacts (active accounts only)	HubSpot Contacts	~1,000-2,000	Contact	Contacts on migrated accounts that weren't in Tier 1
Lead Source data	HubSpot Contact properties	~small subset	Contact/Lead	Only where Lead Source is populated (4.7% fill rate)

Lead Object Migration Dependency

Note (v1.1): Lead migration strategy depends on the resolution of Open Question Q23 (Lead vs Contact Architecture Decision). If the contacts-only approach is adopted after sandbox experimentation, all HubSpot contacts will migrate as Contacts (no Lead object population from migration). This decision will be made during Phase 1.

What Does NOT Migrate

Data Set	Reason	Volume Left Behind
Stalled Qualify deals	Ghost pipeline — 81.9% stalled, never worked	~158 deals
Contacts with no account association	Unowned, unassociated — no value in SF	~51.5% of contacts
Deals with \$0 amount	No revenue data — nothing to report on	Most HubSpot deals
95%-empty custom properties	Would create empty fields in SF	59/62 contact properties
Disabled workflow history	Broken automation artifacts	6 workflows
Smart Maintenance Pilots pipeline	0 deals — never used	Empty pipeline
Product C Contract Conversion pipeline	0 deals — never used	Empty pipeline
NARPM convention contacts (pending Q16)	Potentially stale at 9+ months	~2,632 contacts
Duplicate contacts	Dedup during staging, not after import	TBD volume

Field Mapping: HubSpot to Salesforce

Company > Account

HubSpot Property	SF Field	Transformation
name	Name	Direct
domain	Website	Direct
phone	Phone	Direct
state	BillingState	Convert to full state name if abbreviated
city	BillingCity	Direct
industry	Industry	Map to SF standard picklist values
lifecyclestage	Account_Status__c	Map: customer > Active , opportunity > Opportunity Active , all others > Prospect
N/A	RecordTypeId	Set based on lifecycle stage
hubspot_owner_id	OwnerId	Map HubSpot owner email > SF User ID

Contact > Contact

HubSpot Property	SF Field	Transformation
firstname	FirstName	Direct
lastname	LastName	Direct
email	Email	Direct — primary dedup key
phone	Phone	Direct
jobtitle	Title	Direct
associatedcompanyid	AccountId	Map via Company migration — reject if no Account match
hubspot_owner_id	OwnerId	Map HubSpot owner email > SF User ID. If null, assign to Account owner.
hs_lead_status	Prospect_Status__c	Map: see Demand Funnel mapping below
N/A	Demand_Funnel_Stage__c	Derived from deal history: see Demand Funnel mapping below

Deal > Opportunity

HubSpot Property	SF Field	Transformation
dealname	Name	Direct
amount	Amount	Direct — reject if 0 or null for Closed Won
closedate	CloseDate	Direct
dealstage	StageName	Map HubSpot stage IDs to SF stage names per Record Type
pipeline	RecordTypeId	Map HubSpot pipeline > SF Record Type
dealtype	Record Type assist	Use to determine New Business vs Amendment vs Renewal
hubspot_owner_id	OwnerId	Map HubSpot owner email > SF User ID
associatedcompanyid	AccountId	Map via Company migration
N/A	Closed_Lost_Reason__c	Map from HubSpot closed lost reason if available

Demand Funnel Stage Mapping (Migration)

All migrated contacts receive a `Demand_Funnel_Stage__c` value based on their deal history:

Condition	Demand Funnel Stage
Contact on Account with Closed Won Opportunity	Customer
Contact on Account with open Opportunity (Stage 1+)	SQO
Contact on Account with open Opportunity (Stage 0)	SAL
Contact with <code>Recent_Conversion_Date__c</code> populated	SQL
Contact on active Account with no Opportunity	Engaged
All other migrated contacts	Suspect

Prospect Status Mapping (Migration)

All migrated contacts receive `Prospect_Status__c` = `Open` (fresh start for sales engagement tracking). Post-migration, Prospect Status will be driven by sales activity and automation.

Pipeline > Record Type Mapping

HubSpot Pipeline	SF Record Type	Notes
Core New Business	New Business	—
Core Expansion	Amendment	—
Core Renewal	Renewal	—
Product B	New Business	Product_Type__c = Product B
Product C	New Business	Product_Type__c = Product C
Product C Expansion	Amendment	Product_Type__c = Product C
Product C Renewal	Renewal	Product_Type__c = Product C
Smart Maintenance Pilots	SKIP	0 deals — do not migrate
Product C Contract Conversion	SKIP	0 deals — do not migrate

Migration Process

Step 1: Export from HubSpot (RevPal-managed)

RevPal handles all extraction via HubSpot API and programmatic tools.

1. Export Companies (active customers + active pipeline accounts)
 - Filter: lifecycle_stage IN (customer, opportunity) OR has open deals with activity in last 90 days
 - Fields: name, domain, phone, state, city, industry, lifecyclestage, hubspot_owner_id, hs_object_id
2. Export Contacts (on migrated accounts only)
 - Filter: associatedcompanyid IN (migrated company IDs) AND email IS NOT NULL
 - Fields: firstname, lastname, email, phone, jobtitle, associatedcompanyid, hubspot_owner_id, hs_lead_status
3. Export Deals (actively worked + closed won last 12 months)
 - Filter: (dealstage NOT IN stalled_qualify_stages AND last_activity > 90 days ago) OR (is_closed_won AND closedate > 12 months ago AND amount > 0)
 - Fields: dealname, amount, closedate, dealstage, pipeline, dealtype, hubspot_owner_id, associatedcompanyid

Step 2: Staging & Cleaning

1. Dedup Companies by domain
2. Dedup Contacts by email
3. Validate all owner IDs map to active SF Users – fallback to RevOps user for unmapped
4. Validate all Contacts have a matching Account in the Company export
5. Validate all Deals have a matching Account in the Company export
6. Convert state abbreviations to full names
7. Map HubSpot pipeline/stage IDs to SF Record Type/Stage
8. Remove deals with \$0 amount from Closed Won set
9. Generate CSVs with explicit LF line endings (WSL2 requirement)
10. Produce migration summary: counts by object, by Record Type, by stage

Step 3: Test Migration (Sandbox)

1. Import Accounts (Customer RT first, then Prospect RT)
2. Import Contacts (with Account lookup)
3. Import Opportunities (with Account lookup, correct Record Type)
4. Validate counts match staging summary
5. Spot-check 10 records per object for data integrity
6. Run all SF reports against migrated data – verify counts
7. Walk through UAT scenarios per role

Step 4: Production Migration

1. Freeze HubSpot deal updates (announce cutover window)
2. Re-export any deals updated since staging export
3. Run production import in same order: Accounts > Contacts > Opportunities
4. Validate counts and spot-check
5. Enable HubSpot <> SF sync (post-migration, for ongoing marketing intake)
6. Announce go-live

Owner Mapping

HubSpot Owner	SF User	Role
TBD — need HubSpot owner list	TBD — need SF user list	Map by email match

Action: Export HubSpot owners (`GET /crm/v3/owners`) and SF users (`SELECT Id, Email, Name FROM User WHERE IsActive = true`) to build the mapping table before migration.

Rollback Plan

If the production migration fails or data integrity issues are discovered:

1. **Within 24 hours:** Delete all imported records by batch (tagged with `Migration_Batch__c` field on each object)
2. **Preserve existing SF data:** The org is near-empty (~100 sample records), so collision risk is minimal
3. **Re-run from staging:** Fix the issue in the staging CSV and re-import
4. **HubSpot remains available:** Until formal deactivation, HubSpot can serve as the fallback system

CSV Technical Requirements

All CSVs generated for Salesforce import must use explicit LF line endings:

```
# CORRECT – explicit LF for Salesforce Bulk API on WSL2
with open('file.csv', 'w', newline='\n') as f:
    writer = csv.writer(f, lineterminator='\n')
```

Or post-process: `sed -i 's/\r$//' file.csv`

Timeline

Step	Week	Duration	Dependency
Export from HubSpot	Week 5	1 day	HubSpot access
Staging & Cleaning	Week 5	2-3 days	Export complete
Test Migration (Sandbox)	Week 6	2 days	Sandbox build complete (Phases 0-2)
UAT on migrated data	Week 6-7	3-5 days	Test migration passes
Production Migration	Week 8	1 day	UAT passes + go-live gates met
Post-migration validation	Week 8	1 day	Production import complete