

# **Primary Datasets & API Endpoints (Production-Grade)**

Use these as the backbone. Each row = one feed family to ingest + backfill.

# A1) Conflict, Events, Maritime, Access, Sanctions

What	Why it matters	Where to get it	How to fetch / API
Armed conflict & political violence ( <a href="#">ACLED</a> )	Geocoded events, actors, fatalities; best-in-class for conflict analytics	<a href="#">ACLED</a> (licence/API required)	API documented by <a href="#">ACLED</a> (requires key).
Air war & civilian harm ( <a href="#">Yemen Data Project</a> )	Airstrike-level data since 2015, methodology notes	<a href="#">Yemen Data Project</a> (official site)	(Public site; download pages vary. Treat as periodic bulk imports; validate against <a href="#">ACLED</a> where possible.)
UN Verification & Inspection Mechanism - <a href="#">UNVIM</a> (Hudaydah, Saleef)	Actual ships cleared, by cargo (food, fuel), by month—directly links to food/fuel availability	<a href="#">UNVIM</a> (UN OCHA)	Monthly updates (PDF/HTML). Use scraping + human verification on month close.
Access constraints & humanitarian ops status ( <a href="#">ReliefWeb</a> /OCHA)	Provides formal access notices, security incidents, and ops updates	<a href="#">ReliefWeb - Yemen country page</a>	RSS/JSON feeds per collection (see Section B).
UN Panel of Experts on Yemen (UNSC 2140)	Gold-standard investigative reporting: arms flows, finance networks, sanctionable conduct	<a href="#">UN Documents</a> (UNDOCs)	Annual report; cite doc symbol (e.g., S/2024/731). Scrape index for PDFs/annexes.
Sanctions lists (US/EU/UK/UN)	Legal lists used in banking filters, KYC, compliance, risk	US: <a href="#">OFAC Sanctions List Search</a> (UI & bulk); EU: <a href="#">EU consolidated list</a> ; UK: <a href="#">OFSI consolidated list</a>	US: SLS search + bulk; EU: data.europa.eu bulk; UK: OFSI list portal. Integrate daily.

## Code-style fetch pointers (examples):

```
# ACLED (requires token; example only)
https://api.acleddata.com/acled/read?country=Yemen&event_date=2011-01-01|2050-01-01&limit=0&key=YOUR_KEY
```

```
# UNVIM – monthly reports (scrape + parse)
https://www.vimye.org/ # discover monthly pages
```

```
# OFAC – daily consolidated exports
https://ofac.treasury.gov/sanctions-list-search # UI
https://ofac.treasury.gov/sanctions-lists/consolidated-sanctions-list-data-files # bulk
```

```
# EU consolidated list (JSON/XML, updated daily)
https://data.europa.eu/data/datasets/consolidated-list-of-persons-groups-and-entities-subject-to-eu-financial-sanctions
```

# A2) Humanitarian, Population, Displacement, Food Security

What	Why it matters	Where to get it	How to fetch / API
Funding flows (OCHA FTS)	Ground truth for who-funded-what-by-when; critical for disbursements	<a href="#">FTS API (HPC Tools)</a>	GET /v2/public/fts/flow?countryISO3=YEM (paginate); also orgs, plans.
Displacement (IOM DTM)	IDP numbers, movements, returnees, site-level insights	<a href="#">IOM-DTM Yemen (HDX)</a>	Use CKAN API to programmatically download CSVs per round.
Refugees/Asylum (UNHCR)	Cross-border displacement; caseloads by status	<a href="#">UNHCR Operational Data Portal (Yemen)</a>	Scrape series tables / use UNHCR data services where available. (Primary portal reference.)
Food prices & minimum food basket (WFP VAM)	Local market signals; underpins your FBF indicator	<a href="#">HDX dataset "WFP food prices for Yemen"</a>	CKAN endpoint: .../api/3/action/package_show?id=wfp-food-prices-for-yemen (get resources, then CSV/ZIP).
IPC / FEWS NET	Early warning on food security phases; narratives & maps	<a href="#">FEWS NET (country page)</a>	Use page/RSS scraping for monthly/quarterly outlooks (FEWS APIs limited). (Country portal reference).
WHO health events (e.g., cholera)	Epidemics affecting labour, prices, access	<a href="#">WHO EMRO &amp; WHO HQ</a>	Pull situation reports; store as events + indicator weekly series. (Reference: WHO Pages).
Population baselines (COD/HDX)	Denominators for rates; district/city population	<a href="#">HDX Yemen COD Population</a>	CKAN endpoint for COD packages; version-control snapshots.

## CKAN pattern (HDX) – programmatic:

# Find a package and get its resource file URLs

```
https://data.humdata.org/api/3/action/package_show?id=wfp-food-prices-for-yemen
```

# Then download the resource 'url' (CSV/ZIP). Repeat for COD-population, DTM, market monitors.

# A3) Economy, Prices, FX, Oil, Trade/Ports

What	Why	Where	How / API
World Bank WDI (GDP, poverty, external balance, etc.)	Long, clean macro series back to 1990	<a href="#"><u>World Bank API</u></a>	Example: <a href="https://api.worldbank.org/v2/country/YEM/indicator/NY.GDP.MKTP.CD?format=json&amp;per_page=200">https://api.worldbank.org/v2/country/YEM/indicator/NY.GDP.MKTP.CD?format=json&amp;per_page=200</a>
IMF IFS / SDMX	Monetary & financial: reserves, M2, CPI (where available)	<a href="#"><u>IMF Data Services (SDMX-JSON)</u></a>	<a href="#">.../REST/SDMX_JSON.svc/CompactData/IFS/YEM.FI...</a> (choose series)
Oil - production, exports	Primary FX earner pre-war; still a key scenario driver	<a href="#"><u>U.S. EIA (international data)</u></a>	EIA tables include Yemen series where available (see crude trade tables; production where present).
JODI Oil	Official energy data (requires profile)	<a href="#"><u>JODI Data Portal</u></a>	Use CSV downloads post-auth; store metadata for audit. (Portal reference)
Ports/imports logistics	Ships, tonnage, food/fuel inflow	<a href="#"><u>UNVIM + Logistics Cluster (WFP)</u></a>	Merge UNVIM monthly with WFP Logistics briefs ( <a href="#"><u>ReliefWeb collections</u></a> ).
Exchange rate (dual markets)	Aden vs. Sana'a FX divergence drives prices	<a href="#"><u>Central bank bulletins</u></a> + trusted market monitors	Store both official/bulletin and observed/market quotes; treat as two series. (See <a href="#"><u>CBY</u></a> and market monitors; combine with <a href="#"><u>WFP price sheets</u></a> .)

- For FX and CPI:** Yemen official CPI is sparse/contested post-2015. Use: (1) WFP market-price-based inflation proxy; (2) IMF/UNDP estimates; (3) [CBY bulletins](#). Triangulate (rules in Section D).

# A4) Law, Regulation, Institutions & Must-Have Report Collections

## Law, regulation, institutions (for your "Regulations & Standards" section)

- [Central Bank of Yemen - Aden](#) (official decisions, circulars), [Central Bank - Sana'a](#) (bulletins), [Ministry of Finance](#) (budget docs), [Tax/Customs Authorities](#), [CSO](#) (statistical yearbooks where available).
- Note: official domains and archives shift; your crawler should harvest all available PDFs and persist immutable copies with provenance metadata.
- [UN Sanctions regimes](#) ([UNSC Yemen sanctions list](#) & [Panel reports](#)), [AML/CFT laws](#) (Yemen gazette scans), Payments/E-money regulations (CBY circulars).

(Direct official portals can change; anchor to [ReliefWeb "Documents"](#) + [UNDOCs](#) for persistence; see Section B collection links.)

## B) Must-have "all-years" report collections (download from originals)

Use these collection pages to backfill 2011-2025 and keep current. Store every PDF with metadata (title, date, source, URL, SHA256).

### B1) Humanitarian planning & flows

- [Humanitarian Response Plans \(HRP\) - Yemen](#) (2015-present), [Humanitarian Needs Overviews \(HNO\)](#) - canonical annual baselines and targets. Fetch all years from [ReliefWeb's Yemen collections](#).
- [OCHA Financial Tracking Service \(FTS\)](#) - plan documents + API flows for funding to Yemen.

### B2) Conflict, sanctions, compliance

- [UNSC Panel of Experts on Yemen](#) (annual reports) - evidence base for sanctions, arms networks, revenue streams. Cite doc symbols and download from [UNDOCs](#) (e.g., S/2024/731).
- [OFAC/EU/UK](#) consolidated sanctions lists - daily bulk files for name screening and your SII (Sanctions Intensity Index).

### B3) Macro & sector

- [World Bank Yemen country publications](#) (Economic Updates, Macro Poverty Outlook, project docs). (Use [WB publications portal](#) + project pages.)
- [IMF country pages](#) ([IFS data](#), press releases, staff statements where applicable).
- [UNDP & UN agencies](#) (e.g., [UNDP Assessing the Impact of War on Development in Yemen](#), [WHO health situation reports](#), [WFP Market Monitors](#)). (Use [agency libraries](#); [ReliefWeb aggregates](#).)

### B4) Displacement & population

- [IOM-DTM Yemen](#) - all rounds, methodologies, rapid assessments ([HDX packages](#)).
- [UNHCR Yemen](#) - caseload dashboards and country updates.
- [COD Population](#) - all revisions via [HDX COD portal](#) (for district/governorate).

### B5) Prices, food security, ports

- [WFP VAM - Yemen food prices](#) (by market/commodity), [minimum food basket methodology \(HDX dataset\)](#).
- [UNVIM](#) - monthly vessel/cargo updates (food, fuel).

# C) Concrete API & Download Cheat-Sheet

Put these in etl/jobs/\*.ts and schedule with retry + checksum.

```
# World Bank – GDP current US$ (back to 1960); add per_page to cover all years  
https://api.worldbank.org/v2/country/YEM/indicator/NY.GDP.MKTP.CD?format=json&per_page=200
```

```
# OCHA FTS – all funding flows to Yemen (paginate)  
https://api.hpc.tools/v2/public/fts/flow?countryISO3=YEM
```

```
# IMF IFS – example (reserves, money, prices) via SDMX JSON (choose series IDs)  
http://dataservices.imf.org/REST/SDMX_JSON.svc/CompactData/IFS/YEM.NNOB_XDC.A
```

```
# HDX (CKAN) – get the file URLs (WFP prices)  
https://data.humdata.org/api/3/action/package_show?id=wfp-food-prices-for-yemen
```

```
# ACLED – events (requires token; sample shape)  
https://api.acleddata.com/acled/read?country=Yemen&event_date=2011-01-01|2050-01-01&limit=0&key=YOUR_KEY
```

```
# UNVIM – monthly report list (crawl)  
https://www.vimye.org/
```

```
# EU consolidated sanctions list (bundle)  
https://data.europa.eu/data/datasets/consolidated-list-of-persons-groups-and-entities-subject-to-eu-financial-sanctions
```

```
# OFAC – consolidated sanctions list bulk files  
https://ofac.treasury.gov/sanctions-lists/consolidated-sanctions-list-data-files
```

(See citations above for each family.)

# D) Triangulation & Validation Rules

Always store: raw value, unit, method, vintage, source URL, SHA256, confidence.

Indicator	Primary	Cross-checks	Recon rule (pseudo)
GDP (current US\$)	<a href="#">WB WDI</a>	<a href="#">IMF WEO/IFS</a> narrative	Prefer WB unless WEO newer; if delta > 10% same year, flag for analyst review.
Inflation (proxy)	<a href="#">WFP price indices</a>	<a href="#">IMF CPI</a> , ad-hoc CSO	Compute Laspeyres from WFP commodity basket; backfill with IMF CPI if available; mark as derived.
FX (Aden, Sana'a)	<a href="#">CBY bulletins</a> + market sheets	<a href="#">WFP market prices</a> (implied FX), remittance dealers	Maintain FX_ADEN and FX_SANA; if gap > 15% for > 14 days ⇒ trigger Currency Stress Alert.
Food basket cost	<a href="#">WFP VAM (FBF)</a>	<a href="#">FEWS NET</a> narratives, market monitors	Use WFP; when missing, compute from WFP commodity lines (rice/wheat/fuel) with standard basket weights.
Imports (fuel/food)	<a href="#">UNVIM</a>	<a href="#">WFP Logistics Cluster</a>	Join by month; if UNVIM < 70% of 12-mo avg and FX gap rising ⇒ Supply Shock Alert.
Displacement	<a href="#">IOM DTM</a>	<a href="#">UNHCR</a> , <a href="#">OCHA sitreps</a>	Use DTM rounds; interpolate linearly; major step changes must coincide with <a href="#">conflict event clusters (ACLED)</a> .
Sanctions intensity (SII)	<a href="#">OFAC/EU/UK</a> counts	<a href="#">UN Panel of Experts</a>	SII = weighted(# new listings, severity, domain [finance/arms/shipping]) over 90 days; normalise 0-100.

# E) Database Schema & F) Alerting & Intelligence

## E) Database schema (production)

- actors(id, name, type, area\_of\_influence, aliases[], sanction\_ids[], ...)
- events(id, date, geo, actor\_ids[], type, fatalities, source\_url, source\_sha, confidence)
- indicators(code, name, unit, frequency, description, source\_primary, source\_alt[])
- indicator\_values(code, period, value, method, vintage\_date, source\_url, sha, confidence)
- ports\_imports(month, port, cargo\_type, tonnage, vessels, source\_url, sha)
- sanctions\_entities(id, program, name, aka[], listed\_on, measures[], source\_list, raw\_json)
- projects(id, donor, implementer, sector, amount, start, end, status, docs[])
- regulations(id, issuer, date, subject, summary, link, sensitivity\_tag)
- causations(event\_id ⇒ indicator\_code, lag\_days, effect\_estimate, method, notes)

## F) Alerting & intelligence (how it actually works)

### F1) Signal detectors (streaming jobs)

- **Conflict spike detector:** rolling 7-day ACLED counts per governorate; z-score > 2.0 ⇒ Conflict Surge.
- **Currency stress:** gap% = (FX\_ADEN-FX\_SANA)/mid. If gap% > 15% for 14 days or |d/dt| > 5%/wk ⇒ Currency Fragmentation.
- **Supply shock:** UNVIM fuel+food tonnage < 70% of trailing 12-mo avg AND concurrent Red Sea incident ⇒ Port Supply Risk.
- **Sanctions pulse:** SII 20-point jump in 30 days, or designation of banking/port/shipping nodes ⇒ Compliance Heat.
- **Humanitarian underfunding:** FTS plan coverage < 35% by mid-year ⇒ Funding Gap Critical.

### F2) Causation graph (event → indicator)

Store event-impact edges with lag windows (e.g., Red Sea attack → imports↓ (t+1..t+3) → fuel price↑ (t+1..t+4) → FBF↑ (t+2..t+5)). Train elastic-net or Bayesian structural time series on historicals; keep features interpretable; show confidence.

### F3) Confidence model (0-1)

- Source quality (UN=0.95, WB/IMF=0.9, ACLED=0.85, media tiered 0.5-0.8).
- Cross-confirmation (n distinct top-tier sources) and consistency (no contradiction with prior).
- Recency & completeness (freshness decay; missing fields penalise).
- Final: conf = wqSq + wcCc + wrR + wmCompleteness. Publish with Low/Medium/High badges.

# F4) Auto-Briefs, Distribution & Governance

## F4) Auto-briefs & distribution

### Products

- Flash Alert ( $\leq$  250 words + 1 chart)
- Executive Readout (1-2 pp)
- Weekly Compass (KPI deck)

### Delivery

Email + dashboard + API; Arabic & English side-by-side; "Do No Harm" redlines (no sensitive operational details).

### Branding & disclaimers

Watermark "System-generated; human-reviewed" + confidence.

### Personalisation

e.g., CBY-Aden gets monetary policy read; port authorities get import alerts; WFP gets food/fuel nexus charts.

### Routing

Tag every alert with topic (currency, fuel, ports, food, conflict, sanctions) and stakeholders (donors, CBY, banks, humanitarian ops, press).

### Template (exec):

**Subject:** Yemen | Currency Fragmentation Alert – FX gap rose to 17% (High confidence)

Why it matters (3 bullets) ...

What just happened (data with sources) ...

Likely impacts (ports, fuel, FBF) ...

Recommended actions (by audience) ...

Sources: [FTS API](#), [UNVIM](#), [WFP VAM](#), [ACLED](#) (URLs)

Confidence: High (0.82)

## G) Governance & quality

- **Provenance:** store URL + SHA256 of every file; never overwrite raw.
- **Versioning:** indicators are vintaged (revisions tracked).
- **Peer review:** analyst must approve any High-impact alert before mailout.
- **Bilingual parity:** every chart caption and tooltip translated; glossary tooltips site-wide.
- **Sensitivity tags:** PUBLIC | LIMITED | SENSITIVE-OPS. HRP/HNO are public; certain CBY memos may be LIMITED.

# H) Glossary, I) Sprint Plan & K) Intelligence Layer Example

## H) Glossary (public-facing, embedded tooltips)

- **FX**

Foreign exchange; here, YER per USD (two markets since 2019: Aden vs. Sana'a).

- **FBF (Minimum Food Basket)**

Monthly cost for a standard family's essential diet.

- **FTS**

[OCHA Financial Tracking Service](#)—official registry of humanitarian funding.

- **UNVIM**

[UN mechanism](#) clearing commercial imports to Hudaydah/Saleef during conflict.

- **IPC**

[Integrated Food Security Phase Classification](#) (Phase 1-5).

- **DTM**

[IOM Displacement Tracking Matrix](#) (IDP/returnee assessments).

- **SII**

Sanctions Intensity Index—composite we compute from new listings & scope.

- **CBY**

[Central Bank of Yemen](#) (two authorities since 2016: Aden & Sana'a).

- **ACLED**

[Armed Conflict Location & Event Data](#) (geocoded conflict events).

- **HDX/CKAN**

[Humanitarian Data Exchange](#) and its [API](#) for datasets.

## I) What to ingest first (10-day sprint plan)

01

### Day 1-2 (plumbing)

Stand up secrets, schedulers, S3, SHA256, vintaging tables.

02

### Day 3-4 (feeds)

[WB](#) GDP + [FTS](#) + [WFP prices](#) + [UNVIM](#) + [IOM DTM](#) (last 10 years).

03

### Day 5 (validators)

Implement rules in Section D; create "Data Quality" dashboard.

04

### Day 6-7 (alerts)

Currency, supply, sanctions pulses + email templates + Arabic/English.

05

### Day 8 (visuals)

Master Compass tiles (GDP, FTS coverage %, FX gap, FBF, imports).

06

### Day 9-10 (reports)

Auto-generate Weekly Compass and Flash Alert.

## K) How the "intelligent" layer surfaces the story (worked example)

**Scenario:** Houthi-claimed drone strike on a Red Sea vessel + sharp Aden/Sana'a FX divergence.

**System chain:**

1. Event ingested ([ACLED](#) + [Reuters](#)/[UNVIM](#) page) ⇒ conflict spike in Red Sea grid cells.
2. Imports forecast (state-space model): next 1-3 months fuel/food tonnage down 20-30% (CI 60-75%).
3. Price pass-through (basket model): FBF up +8-12% lagging 1-2 months.
4. Funding check ([FTS](#)): HRP fuel logistics funding at 24% (below threshold).
5. Stakeholder-specific alerts:
  - [CBY](#)-Aden: FX gap → consider temporary FX auctions for fuel importers; publish circular.
  - [WFP](#)/Logistics Cluster: prioritise fuel pipelines; adjust cash transfer values.
  - Donors: fast-track logistics line items.

6. Auto-brief generated (AR/EN), confidence High (0.78), attached chart pack (imports, FX gap, FBF).

## Final notes for the team

- Build incremental crawlers for [ReliefWeb](#), [UNDOCS](#), [UNVIM](#), [HDX \(CKAN\)](#), [FTS](#), [OFAC/EU/UK](#), [EIA](#).
- Keep a single Source Registry table with: owner, endpoint, cadence, auth, last\_success, last\_hash, licence.
- Every number on-site is footnoted: Source name + link + vintage.
- Bilingual content by default; glossary tooltips everywhere.
- "Do No Harm" is enforced in code: the mailer refuses to send anything with SENSITIVE-OPS.