Source structure and extraction atlas

Companion to *Organological databases and repositories: a critical source catalogue*

Dominik Ukolov | 11 September 2026 | Unpublished

Prepared for reserved DOI 10.5281/zenodo.22693083. This companion describes the same 61 catalogue entries as the study: 53 registered source keys, five subordinate score collections and three audit-only resources. The entries are not 61 independent organ databases.

Purpose and evidence

The atlas explains how documentary objects are arranged, how Aggregator selects and interprets them, and which distinctions are lost or retained. It combines the dated source accounts and selected-document observations in the study with inspection of Aggregator revision `2bb6e85282e3ab11af1bcc7126290f67af176f53`. Source observations remain dated 10-11 September 2026; inspecting the code on 11 September does not constitute a new visit to every website. The source catalogue supplies the inclusion frame, attribution and original references.[1]

The consistent chapter structure separates identity, native contents, discovery, transformation, historical relationships, loss and uncertainty, implementation evidence, and research use. This organisation draws on dataset documentation's attention to composition, collection, processing and use, and on provenance's distinction between entities, activities and responsible agents. It is an editorial application of those ideas, not a claim of formal conformance or certification.[2,3]

Evidence vocabulary. Source account means a description supported by the cited publisher page. Selected-document observation means a statement about the particular document examined, with its date and reference. Code reading means an observation of the pinned implementation. Interpretation means a reasoned consequence of those materials. These levels are not interchangeable. A coded branch is not proof that it ran; a populated output is not proof that the underlying fact is correct; an empty result may reflect failed acquisition, absent input or an unsupported layout.

Coverage. Every entry in the selection frame receives the same eight questions. Where no registered adapter or inspected detail document exists, the corresponding gap is stated. This is complete coverage of the defined catalogue and the stated static-inspection procedure. It is not a census of all source pages, a reverse-engineered upstream database schema, or a validation of every runtime branch. No new crawl, download campaign, database import or full-corpus replay was performed for this atlas.

Implementation method. Python source was parsed without importing the application. Effective class methods were resolved through locally declared inheritance, with later same-class definitions taking precedence. The machine-readable register records selected domain and acquisition methods, literal dictionary reads, HTML queries and dictionary constructions. These are syntactic observations with exact file and line locators. They are not automatically inferred field mappings. Editorial readings trace the important source-specific behaviours; the inherited static HTML family and imported normalizers are described separately. Runtime configuration, external dependencies and later Processor operations are outside this static inventory.

Reproducibility and citation. Cite the atlas by catalogue ID and subsection, and use the JSON record for its code revision, file hash, function and expression. Native claims retain the original bibliography IDs and consulted URLs. A test of the pinned public GitHub locator returned 404 during preparation; independent public access to the repository is therefore not established. The deposited register preserves the inspected expressions and bounded code passages, but not a complete executable repository. A hash identifies bytes; it cannot replace unavailable source documents. The atlas makes a reproducible inspection record available within those limits.

Editorial responsibility. OpenAI Codex assisted code inspection, extraction and drafting. Native observations are inherited from the cited study; implementation findings are a static analysis of the stated revision. No independent peer review, physical instrument examination or extraction-accuracy estimate is claimed. Original field labels, source names and bibliographic titles are retained where translation would obscure exact identification; explanatory prose is English.

How to read the technical evidence

Each chapter identifies its registered class and applicable parser family. A short rule table records selected dictionary constructions or input reads from its transform or source-specific parser. The expression is shown as written semantically by Python's syntax parser; whitespace and quoting may be normalized. A row such as `stop_count = sum(...)` documents a calculation, whereas `orgel.get('anzpfeifen')` documents a field read. A dictionary key can occur in a temporary object: the row is not an assertion of a final JSON path. Full eligible expressions, queries and method locators are in `source_atlas.json`.

The common pipe-organ normalizer projects identity, location, events, participants, specifications, media, references, descriptions, source snapshots and extensions. These are MODAVIS categories, not evidence that every source has those categories natively. Static HTML subclasses can inherit a first-year fallback and a construction-event projection; source-specific overrides sometimes correct that behaviour. The reader must distinguish an inherited heuristic from an explicit source statement.[4,5]

Product sources instead describe catalogue items, versions, assets and dependencies; forum adapters describe threads or posts; score collections describe works and editions/files; iconography describes artworks and depictions. Applying the pipe-organ schema's field checklist to those families would conceal their actual documentary units. Accordingly the atlas supplies common questions rather than binary feature scores.

Source index

S001. Aeolian-Skinner Archive 6

S002. Arp Schnitger Organs 8

S003. Augustine's Virtual Organs Sample Set Listings 10

S004. BDO Tuning Systems 12

S005. Binaural Pipes Instruments 14

S006. Contrebombarde Sample Set Library 16

S007. Digital Sheet Music Sources 18

S008. Estey Pipe Organ Archive 20

S009. Evensong Hauptwerk Sample Sets 22

S010. GDO Literaturdatenbank 24

S011. GdO Orgelbauer Links 26

S012. GrandOrgue GitHub Wiki: Sample Sets 28

S013. INdO 30

S014. IOHIO 32

S015. Jeux d'orgues Sample Set Downloads 34

S016. Kirchenorgel-Forum 36

S017. Lars Palo Free GrandOrgue Samplesets 38

S018. Lavender Audio Hereford Hauptwerk Sample Sets 40

S019. Milan Digital Audio Instruments 42

S020. MPS Orgelseite Filebase 44

S021. Musique et musiciens - Orgues 46

S022. musiXplora 48

S023. Nomine 50

S024. NPOR 52

S025. NYC AGO 54

S026. OHS Pipe Organ Database 56

S027. OHTA Organs of Australia 58

S028. Organ Index 60

S029. Orgbase 62

S030. Orgel-Archiv Schmidt 64

S031. Orgeldatenbank Bayern 66

S032. Orgeldokumentationszentrum HSLU 68

S033. Orgeldatenbank Sachsen 70

S034. Die Orgelseite 72

S035. Orgelsite NL 74

S036. Orgelverzeichnis Schweiz / Liechtenstein 76

S037. OrguesCH 78

S038. Orgues France 80

S039. OSWA Pipe Organs in Western Australia 82

S040. PCorgan Hauptwerk Sample Set Registry 84

S041. Piotr Grabowski GrandOrgue Versions 86

S042. Piotr Grabowski Hauptwerk Instruments 88

S043. Pipedreams Organ Gallery 90

S044. Pipe Organ Map 92

S045. RIdIM Music Iconography 94

S046. Sakralorgelforum 96

S047. SONORA 98

S048. Sonus Paradisi Free and Demo Sample Sets 100

S049. SourceForge GrandOrgue Sample Sets 102

S050. Varhany 104

S051. VI-Control 106

S052. Virtual Pipe Organs Free Sample Sets and Producers 108

S053. Virtual Pipe Organs ODFs For GrandOrgue 110

S054. PDMX Public Domain MusicXML 112

S055. Little Organ Book / Orgelbuechlein Dataset 114

S056. IMSLP / Werner Icking Music Collection 116

S057. Mutopia Project 118

S058. MuseScore.com Direct 120

S059. ORKASA 122

S060. Encyclopedia of Organ Stops 123

S061. Welt-Orgel-Archiv 124

S001. Aeolian-Skinner Archive

Catalogue key: aeolianskinner | Role: registered source | Documentary unit: Factory specification or manufacturer revision

1. Identity, scope and evidence

Source account. The archive presents stoplists, photographs and recollections arranged around the firm's instruments. Specifications describe delivery from the factory and revisions by Skinner or Aeolian-Skinner; subsequent alterations by other builders are excluded.[6,7]

Attribution recorded in the catalogue: Steve Layden (website creator); Jeff Scofield (website manager). Scope: Skinner and Aeolian-Skinner instruments.

2. Native record structure and contents

Selected document; 2026-09-11. Static opus page with identification block, specification and documentary notes[7]

- Town, state and church head the record.
- Named manufacturer and opus number.
- Delivery year and a separately dated later alteration.
- Division-based stops include pitches and borrowing references.
- Separate Great, Swell, Choir, Echo and Pedal groups.
- The later note describes replacement of the console.
- Photograph and named photo supplier.
- A named specification source is supplied.
- Cross-reference to an OHS instrument record.
- Specification and photograph have separate source credits.

Repeated structures: Division columns contain stop name, pitch and pipe quantity or borrowing reference.

3. Discovery and acquisition

The archive subclass has dedicated index, title and specification extraction for opus pages. Its source scope is the manufacturer configuration described by the archive, subject to the limitations stated in the cited source account.[5]

Configured source address: https://archive.pipeorgandatabase.org/aeolianskinner. This is a registry setting, not a new availability test.

Configured primary identifier: aeolianskinner_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "United States" (aggregator/static_html_crawler.py, line 1177).
- start_urls: ["https://archive.pipeorgandatabase.org/aeolianskinner/aeolian_skinner.html", "https://archive.pipeorgandatabase.org/aeolianskinner/skinner.html", "https://archive.pipeorgandatabase.org/aeolianskinner/e_m_skinner_son.html", "https://archive.pipeorgandatabase.org/aeolianskinner/bassett_skinner.html"] (aggregator/static_html_crawler.py, line 1178).
- detail_url_patterns: ["/aeolianskinner/specs/op"] (aggregator/static_html_crawler.py, line 1184).

4. Extraction and normalization

The specialised stoplist routine handles division-oriented specification material, while the shared transform supplies other metadata, raw text, media and source-qualified identity. Pipe quantities and borrowing annotations in the observed page belong to individual stop entries and the stated factory configuration.[5]

5. Dates, relationships and historical states

The specification represents the manufacturer state. Later alterations appear in a separate narrative note.[7]

An opus number can join this archive to an instrument register, but the specifications must be attached to the manufacturer's documented configuration. They cannot stand for the instrument's present state.[6,7]

6. Preservation, loss and uncertainty

The inspected source may also contain later documentary notes. A note about a replacement console does not redate the factory stoplist. Neither the generic builder/year projection nor a matched opus supplies independent evidence of the current configuration.[5]

Limit of the selected-document observation: A stoplist and a later note can concern different states of the same instrument.

7. Implementation evidence and rule register

Registered class: AeolianSkinnerArchiveCrawler in aggregator/static_html_crawler.py; inheritance: AeolianSkinnerArchiveCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- StaticHtmlOrganCrawler.transform (lines 434-516) [inherited].
- AeolianSkinnerArchiveCrawler._extract_raw_spec (lines 1222-1266).
- AeolianSkinnerArchiveCrawler._extract_title (lines 1206-1209).
- AeolianSkinnerArchiveCrawler._location_parts (lines 1211-1220).
- AeolianSkinnerArchiveCrawler._looks_like_aeolian_stop (lines 1269-1271).
- AeolianSkinnerArchiveCrawler._parse_index_page (lines 1186-1204).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	StaticHtmlOrganCrawler.transform, line 464
manual_count	manual_count	StaticHtmlOrganCrawler.transform, line 463
rank_count	rank_count	StaticHtmlOrganCrawler.transform, line 465
stops_by_division	stops_by_division	StaticHtmlOrganCrawler.transform, line 466
country	raw_data.get('country') or self.default_country	StaticHtmlOrganCrawler.transform, line 453
raw_str	title	StaticHtmlOrganCrawler.transform, line 457

8. Research use, citation and rights

The archive displays its own copyright notice; no general open licence was established.[6,7]

S002. Arp Schnitger Organs

Catalogue key: arpschnitger | Role: registered source | Documentary unit: Builder-focused organ profile

1. Identity, scope and evidence

Source account. This personal site documents Schnitger organs and also carries material on organs in Drenthe. Its entry page names Geert Jan Pottjewijd and gives 4 April 2026 as its last update. Historical announcements remain on the same page.[8,9]

Attribution recorded in the catalogue: Geert Jan Pottjewijd. Scope: Arp Schnitger organs; related material on Drenthe.

2. Native record structure and contents

Selected document; 2026-09-11. Static historical essay with photographs and a later specification[9]

- Glückstadt and the subsequent destination of the case are named.
- Several builders are associated with dated interventions.
- Chronological entries distinguish construction, replacement and relocation.
- Later specification arranged by divisions and pitches.
- Hauptwerk, Rückpositiv and Pedal are separate groups.
- Case survival and transfer are described.
- The original organ is identified as former; the case survives elsewhere.
- Historical and later case photographs with credits.
- Named information provider and an image-source citation.
- Links to the later case photograph and its Wikimedia source.
- Information and photograph contributors are named.

Repeated structures: Dated narrative entries and a division-based stop table.

3. Discovery and acquisition

Configured index and detail-link patterns identify the specialist site's organ pages. The subclass supplies page-title and location conventions.[5]

Configured source address: https://arpschnitger.nl. This is a registry setting, not a new availability test.

Configured primary identifier: arpschnitger_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- start_urls: ["https://arpschnitger.nl/schnit.html"] (aggregator/static_html_crawler.py, line 1344).
- detail_url_patterns: ["arpschnitger.nl/s"] (aggregator/static_html_crawler.py, line 1345).
- exclude_url_patterns: ["schnit.html", "schnintro", "sdiscogr", "schnitlit", "schnitdvd", "schnittek", "youtube", ".pdf", ".wmv"] (aggregator/static_html_crawler.py, line 1346).

4. Extraction and normalization

The transform delegates its substantive extraction to the common static-page implementation. The distinction matters: a named source class does not mean that its historical narratives have a dedicated event parser. Raw text, metadata, table material and images are preserved alongside the common summary projection.[5]

5. Dates, relationships and historical states

One page follows a case across replacement instruments and relocation to another church.[9]

The builder-centred selection makes this a useful specialist counterpart to regional inventories. The date on the home page does not date every instrument description, and the different sections should not be counted as one geographical census.[8,9]

6. Preservation, loss and uncertainty

A case can move to another location and become part of another instrument. The observed source account documents precisely this problem. A single URL, current-location label or first year cannot establish uninterrupted instrument identity or one configuration.[5]

Limit of the selected-document observation: A page is not necessarily one instrument at one place. The later stoplist belongs to the organ installed in the transferred case.

7. Implementation evidence and rule register

Registered class: ArpSchnitgerCrawler in aggregator/static_html_crawler.py; inheritance: ArpSchnitgerCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- ArpSchnitgerCrawler.transform (lines 1398-1399).
- ArpSchnitgerCrawler._extract_title (lines 1392-1396).
- ArpSchnitgerCrawler._location_parts (lines 1401-1408).
- ArpSchnitgerCrawler._parse_index_page (lines 1379-1390).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
task_id (read)	kwargs.get('task_id')	StaticHtmlOrganCrawler.crawl, line 79
max_items (read)	kwargs.get('max_items')	StaticHtmlOrganCrawler.crawl, line 80
start_urls (read)	kwargs.get('start_urls')	StaticHtmlOrganCrawler.crawl, line 81
rate_limit_delay (read)	self.config.get('rate_limit_delay', 0.2)	StaticHtmlOrganCrawler.crawl, line 117
timeout (read)	self.config.get('timeout', 30)	ArpSchnitgerCrawler._fetch_html, line 1367

8. Research use, citation and rights

No general reuse licence was established from the consulted entry page.[8,9]

S003. Augustine's Virtual Organs Sample Set Listings

Catalogue key: augustine_virtual_organs_samples | Role: registered source | Documentary unit: Original sample-set listing

1. Identity, scope and evidence

Source account. The catalogue separates original recordings from composite sets and marks a subset as free. It also distinguishes the producer's own recordings from recordings supplied by others and subsequently made into sample sets.[10]

Attribution recorded in the catalogue: Augustine's Virtual Organs. Scope: Hauptwerk products grouped by style, builder and manual count.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects links to PHP product pages and their surrounding context. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://hauptwerk-augustine.info/Sample_sets.php. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_registry" (aggregator/grandorgue_library_crawler.py, line 1402).

4. Extraction and normalization

The implemented unit is sample-set entries. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Original, composite and derived products need separate records. The free category describes distribution price; it does not supply a common licence for the audio.[10]

6. Preservation, loss and uncertainty

Original, composite and derived products need separate provenance. A composite cannot stand as a measured historical organ configuration. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: AugustineVirtualOrgansSamplesCrawler in aggregator/grandorgue_library_crawler.py; inheritance: AugustineVirtualOrgansSamplesCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- AugustineVirtualOrgansSamplesCrawler.parse_html (lines 1404-1466).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'producer_listing'	AugustineVirtualOrgansSamplesCrawler.parse_html, line 1450
requires_sample_set	False	AugustineVirtualOrgansSamplesCrawler.parse_html, line 1452
is_odf_only	False	AugustineVirtualOrgansSamplesCrawler.parse_html, line 1451
producer_label	"Augustine's Virtual Organs"	AugustineVirtualOrgansSamplesCrawler.parse_html, line 1441

8. Research use, citation and rights

Copyright is asserted in the footer; product-specific permissions remain to be checked.[10]

S004. BDO Tuning Systems

Catalogue key: bdo_tuning_systems | Role: registered source | Documentary unit: Tuning-system definition

1. Identity, scope and evidence

Source account. The collection groups temperaments, offers sound examples and distinguishes a more precise calculation section. Its authors page credits Frischknecht with the conception, research and graphical presentation and Schmid with the website's implementation.[13,14]

Attribution recorded in the catalogue: Hans Eugen Frischknecht; Jakob Schmid; hosted by Bund Deutscher Orgelbaumeister. Scope: Historical and theoretical temperaments.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The crawler parses tuning definitions and can obtain separate commentary pages. Source parameter values and commentary have separate document identities.[15,16]

Configured source address: https://bund-deutscher-orgelbaumeister.de/stimmungen/. This is a registry setting, not a new availability test.

Configured primary identifier: tuning_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The normalizer preserves a cent-value vector in a declared source tone order, group memberships and a precise-variant flag. Derived intervals are calculated in a separate routine. Commentary is partitioned into description, author note, source note and literature information. Source values and calculated musical quantities are therefore explicitly distinguishable.[15,16]

5. Dates, relationships and historical states

Temperament definitions support comparison of interval structures. They provide no evidence that a particular organ used a temperament.[13,14]

6. Preservation, loss and uncertainty

A tuning definition does not demonstrate its use on an organ. A precise-variant suffix is an implementation convention that needs source interpretation. Calculated intervals should be cited as calculations with their parameters, not as verbatim source values.[15,16]

7. Implementation evidence and rule register

Registered class: BdoTuningSystemsCrawler in aggregator/bdo_tuning_systems_crawler.py; inheritance: BdoTuningSystemsCrawler > BaseCrawler.[15,16]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- BdoTuningSystemsCrawler.crawl (lines 87-132).
- BdoTuningSystemsCrawler.transform (lines 161-177).
- BdoTuningSystemsCrawler.parse_catalog_page (lines 148-159).
- BdoTuningSystemsCrawler.parse_commentary_page (lines 179-213).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
raw_data_json (read)	payload.get('raw_data_json')	BdoTuningSystemsCrawler.transform, line 163
crawl_run_id (read)	payload.get('crawl_run_id')	BdoTuningSystemsCrawler.transform, line 175
transform_run_id (read)	payload.get('transform_run_id')	BdoTuningSystemsCrawler.transform, line 176
raw_html (read)	payload.get('raw_html')	BdoTuningSystemsCrawler.transform, line 174
task_id (read)	kwargs.get('task_id')	BdoTuningSystemsCrawler.crawl, line 89
limit (read)	options.get('limit')	BdoTuningSystemsCrawler.crawl, line 91

8. Research use, citation and rights

No general licence was identified on the consulted index or authors page.[13,14]

S005. Binaural Pipes Instruments

Catalogue key: binauralpipes_instruments | Role: registered source | Documentary unit: Instrument or prospective sample-set listing

1. Identity, scope and evidence

Source account. The instruments page lists places, builders, dates and compact stop/manual counts. It includes a section for sample sets in preparation and credits Dietrich Wierczeyko in its copyright notice.[17]

Attribution recorded in the catalogue: Dietrich Wierczeyko / Binauralpipes. Scope: Selected organs represented in a sampling project.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects article content and links to instrument pages on the producer host. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://binauralpipes.com/instrumente. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_catalog" (aggregator/grandorgue_library_crawler.py, line 1151).

4. Extraction and normalization

The implemented unit is instrument/product listings. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

A listed organ is not necessarily an available download. Preparation, release and subsequent revision should be distinct states. The project's selection cannot measure the distribution of surviving organs.[17]

6. Preservation, loss and uncertainty

A listed instrument may describe a planned product. Availability and release state must come from the individual item. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: BinauralPipesInstrumentsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: BinauralPipesInstrumentsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- BinauralPipesInstrumentsCrawler.parse_html (lines 1153-1208).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	BinauralPipesInstrumentsCrawler.parse_html, line 1195
manual_count	manual_count	BinauralPipesInstrumentsCrawler.parse_html, line 1193
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'sample_set'	BinauralPipesInstrumentsCrawler.parse_html, line 1184
requires_sample_set	False	BinauralPipesInstrumentsCrawler.parse_html, line 1186
is_odf_only	False	BinauralPipesInstrumentsCrawler.parse_html, line 1185

8. Research use, citation and rights

Copyright is asserted; the consulted list does not establish download or redistribution terms.[17]

S006. Contrebombarde Sample Set Library

Catalogue key: contrebombarde_organs | Role: registered source | Documentary unit: Sample-set registry entry

1. Identity, scope and evidence

Source account. The search interface distinguishes sample-set producer from physical organ builder and offers filters for style, manuals and construction period. The interface therefore distinguishes the virtual product's producer from the instrument builder.[18]

Attribution recorded in the catalogue: Contrebombarde Concert Hall. Scope: Cross-producer virtual organ catalogue associated with a recording community.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects organ detail/view links and organ-card, article or table containers. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://www.contrebombarde.com/concerthall/organs/search. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "catalog" (aggregator/grandorgue_library_crawler.py, line 652).

4. Extraction and normalization

The implemented unit is registry items reached from search and detail pages. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Use the registry for product discovery and links to performances. Availability, price and licensing should be cited to the producer's corresponding release. A recording associated with a set does not document a performance on the physical organ.[18]

6. Preservation, loss and uncertainty

Registry descriptions and producer claims have different authorship; search traversal does not validate product availability. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: ContrebombardeOrgansCrawler in aggregator/grandorgue_library_crawler.py; inheritance: ContrebombardeOrgansCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- ContrebombardeOrgansCrawler.crawl (lines 657-712).
- ContrebombardeOrgansCrawler.parse_html (lines 714-785).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	_first_int(fields.get('stops'))	ContrebombardeOrgansCrawler.parse_html, line 771
manual_count	_first_int(fields.get('manuals'))	ContrebombardeOrgansCrawler.parse_html, line 769
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'catalog_entry'	ContrebombardeOrgansCrawler.parse_html, line 763
producer_label	fields.get('producer')	ContrebombardeOrgansCrawler.parse_html, line 757

8. Research use, citation and rights

Registry access was public; no common licence for the linked products or performances was established.[18]

S007. Digital Sheet Music Sources

Catalogue key: digital_sheet_music | Role: registered source | Documentary unit: Collection-qualified score record and file assets

1. Identity, scope and evidence

Source account. This key denotes a connector spanning five named collections. Its configured MusicXML.com address is a directory of music resources, not the publisher of all the scores. The five collection definitions are recorded separately in this package.[19,20]

Attribution recorded in the catalogue: MODAVIS Aggregator (composite connector). Scope: PDMX; Little Organ Book; IMSLP/WIMA; Mutopia; direct MuseScore records.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

A single registered connector dispatches to five named collections. Dataset imports, web catalogues and direct score routes have different acquisition conditions; the MusicXML directory is not their common publisher.[21,22]

Configured source address: https://www.musicxml.com/music-in-musicxml/. This is a registry setting, not a new availability test.

Configured primary identifier: score_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The score normalizer separates work metadata from downloadable assets, formats, collection identity and licence statements. It preserves collection-specific data. In the PDMX path it suppresses a numeric instrumentation value when that value simply repeats the track count. This prevents a numeric count being relabelled as instrument names.[21,22]

5. Dates, relationships and historical states

Cite the score's collection, edition and file version. Neither a composition title nor the connector key identifies an edition.[19,20]

6. Preservation, loss and uncertainty

Work, arrangement, edition, encoding and downloadable file are different units. A format label does not guarantee that a file was decoded or compared musically. Collection-specific licences cannot be replaced by an umbrella connector licence.[21,22]

7. Implementation evidence and rule register

Registered class: DigitalSheetMusicCrawler in aggregator/digital_sheet_music_crawler.py; inheritance: DigitalSheetMusicCrawler > BaseCrawler.[21,22]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- DigitalSheetMusicCrawler.crawl (lines 145-164).
- DigitalSheetMusicCrawler.transform (lines 166-179).
- DigitalSheetMusicCrawler._extract_tar_gz_member_from_url (lines 780-812).
- DigitalSheetMusicCrawler._parse_mutopia_listing (lines 907-988).

• DigitalSheetMusicCrawler._respectful_pause (lines 1021-1024).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
score_id (read)	raw_data.get('score_id')	DigitalSheetMusicCrawler.transform, line 167
id (read)	raw_data.get('id')	DigitalSheetMusicCrawler.transform, line 167
raw_data_json (read)	payload.get('raw_data_json')	DigitalSheetMusicCrawler.transform, line 169
transform_run_id (read)	raw_data.get('transform_run_id')	DigitalSheetMusicCrawler.transform, line 178
task_id (read)	kwargs.get('task_id')	DigitalSheetMusicCrawler.crawl, line 148
ok (read)	preflight.get('ok')	DigitalSheetMusicCrawler.crawl, line 150

8. Research use, citation and rights

Rights differ by collection, edition and file. The MusicXML directory supplies no umbrella permission.[19,20]

S008. Estey Pipe Organ Archive

Catalogue key: estey | Role: registered source | Documentary unit: Opus-list entry or archival item

1. Identity, scope and evidence

Source account. The virtual museum joins a pipe-organ opus list to catalogues, photographs, facade drawings, patents, recordings and company history. Although now served from the Pipe Organ Database archive host, the page identifies itself as a privately owned website and carries a Matrix Associates copyright notice.[23,24]

Attribution recorded in the catalogue: Matrix Associates / The Estey Organ virtual museum. Scope: Estey company history, instruments and documentary material.

2. Native record structure and contents

Selected document; 2026-09-11. Static opus record combining stop table and transcribed historical notice[24]

- Church, town and state identify the installation.
- Estey opus identity is supplied by the record and archive context.
- The historical notice is dated to the instrument opening.
- Named stops and pitches grouped under Great, Swell and Pedal.
- Pedal compass and division labels are given.
- Tubular-pneumatic action is stated.
- Couplers and swell control are listed.
- The page incorporates a dated historical press account.

Repeated structures: Manual and pedal groups followed by couplers and action details.

3. Discovery and acquisition

The opus index and individual archive pages have separate parsers. Detail titles and specifications receive archive-specific treatment rather than relying entirely on the generic static-page routines.[5]

Configured source address: https://archive.pipeorgandatabase.org/estey. This is a registry setting, not a new availability test.

Configured primary identifier: estey_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "United States" (aggregator/static_html_crawler.py, line 1084).
- start_urls: ["https://archive.pipeorgandatabase.org/estey/opuslist.html"] (aggregator/static_html_crawler.py, line 1085).
- detail_url_patterns: ["/estey/opus"] (aggregator/static_html_crawler.py, line 1086).

4. Extraction and normalization

The specialised specification extractor reads manual/pedal sections of the archive layout; the remainder of the transfer uses the shared location, metadata, media and source-text machinery. The documentary unit should remain the opus/installation account, with its press notices or other historical documents separately attributed.[5]

5. Dates, relationships and historical states

The dated newspaper notice is a documentary account, not an update date for the website.[24]

Cite the original archive and individual item as well as the current host. Manufacturing totals in the introduction are historical company statements, not counts of surviving instruments or digitised records.[23,24]

6. Preservation, loss and uncertainty

A factory stoplist, later alteration and transcribed notice can have different dates. The adapter does not settle continuity between an opus and a surviving instrument. Archive-specific stop extraction does not justify treating all inherited generic event rules as historically validated.[5]

Limit of the selected-document observation: Keep the opus specification separate from the historical notice and its publication date.

7. Implementation evidence and rule register

Registered class: EsteyArchiveCrawler in aggregator/static_html_crawler.py; inheritance: EsteyArchiveCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- StaticHtmlOrganCrawler.transform (lines 434-516) [inherited].
- EsteyArchiveCrawler._extract_raw_spec (lines 1088-1138).
- EsteyArchiveCrawler._extract_title (lines 1160-1163).
- EsteyArchiveCrawler._location_parts (lines 1165-1169).
- EsteyArchiveCrawler._parse_index_page (lines 1140-1158).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	StaticHtmlOrganCrawler.transform, line 464
manual_count	manual_count	StaticHtmlOrganCrawler.transform, line 463
rank_count	rank_count	StaticHtmlOrganCrawler.transform, line 465
stops_by_division	stops_by_division	StaticHtmlOrganCrawler.transform, line 466
country	raw_data.get('country') or self.default_country	StaticHtmlOrganCrawler.transform, line 453
raw_str	title	StaticHtmlOrganCrawler.transform, line 457

8. Research use, citation and rights

The page invites downloading photographs and information and requests attribution when they are reused; it does not name a standard open licence.[23,24]

S009. Evensong Hauptwerk Sample Sets

Catalogue key: evensong_hauptwerk_samplesets | Role: registered source | Documentary unit: Hauptwerk sample-set product

1. Identity, scope and evidence

Source account. The catalogue pairs named organs with product pages and describes the producer's sampling history. It acknowledges extensions using stops from other sources in an early product. A separate sample-set licence is linked from the catalogue.[25,26]

Attribution recorded in the catalogue: Evensong LLC. Scope: Selected instruments, prominently North American organs.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects WordPress page links, captioned figures and linked images. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://www.evensongmusic.net/?page_id=155. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_catalog" (aggregator/grandorgue_library_crawler.py, line 1525).

4. Extraction and normalization

The implemented unit is linked product entries with image/title context. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

A product's virtual specification may exceed its recorded instrument. The package version and the original or extended configuration must therefore accompany comparisons of stops and manuals.[25,26]

6. Preservation, loss and uncertainty

A linked figure can be promotional navigation. Product identity and represented instrument identity require separate evidence. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: EvensongHauptwerkSamplesetsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: EvensongHauptwerkSamplesetsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- EvensongHauptwerkSamplesetsCrawler.parse_html (lines 1527-1651).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'sample_set'	EvensongHauptwerkSamplesetsCrawler.parse_html, line 1556
requires_sample_set	False	EvensongHauptwerkSamplesetsCrawler.parse_html, line 1558
is_odf_only	False	EvensongHauptwerkSamplesetsCrawler.parse_html, line 1557
producer_label	'Evensong'	EvensongHauptwerkSamplesetsCrawler.parse_html, line 1547

8. Research use, citation and rights

A product licence governs use. Its repurposing clause is conditional and should be read for the specific release; no blanket CC0 permission is recorded.[25,26]

S010. GDO Literaturdatenbank

Catalogue key: gdo_literature | Role: registered source | Documentary unit: Bibliographic entry

1. Identity, scope and evidence

Source account. The society credits Alfred and Matthias Reichling with the bibliography. The introduction describes journal surveys, monographs, commemorative volumes, proceedings and articles, with author, title, publication details and annotations. The configured database endpoint returned HTTP 500 during this check; the explanatory page remained readable.[27,28]

Attribution recorded in the catalogue: Alfred Reichling; Matthias Reichling; Gesellschaft der Orgelfreunde. Scope: Literature on organ building and playing; scores excluded.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The crawler acquires bibliographic result rows and associated query context. Cached structured payload and row fields are merged before normalization.[29,30]

Configured source address: https://literaturdb.gdo.de/bib.php. This is a registry setting, not a new availability test.

Configured primary identifier: literature_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The normalizer preserves the citation text and original author label while separately splitting authors, identifying publication years and constructing title, publication-detail and annotation fields. Content-derived identifiers are used when no stable native identifier is supplied. The source page remains the provenance for a bibliographic entry, not the cited publication itself.[29,30]

5. Dates, relationships and historical states

The bibliography is a route to publications, not evidence that their full texts were consulted.[27,28]

6. Preservation, loss and uncertainty

A year extracted from a citation is not necessarily a single publication date; series, volumes and cited historical dates may coexist. Content-hash identity can change with editorial spelling. A catalogue citation is not evidence that the underlying publication was consulted.[29,30]

7. Implementation evidence and rule register

Registered class: GDOLiteratureCrawler in aggregator/gdo_literature_crawler.py; inheritance: GDOLiteratureCrawler > BaseCrawler.[29,30]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GDOLiteratureCrawler.crawl (lines 83-130).
- GDOLiteratureCrawler.transform (lines 187-203).
- GDOLiteratureCrawler._parse_years (lines 467-485).
- GDOLiteratureCrawler.parse_results_html (lines 155-185).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
raw_data_json (read)	payload.get('raw_data_json')	GDOLiteratureCrawler.transform, line 189
crawl_run_id (read)	payload.get('crawl_run_id')	GDOLiteratureCrawler.transform, line 201
transform_run_id (read)	payload.get('transform_run_id')	GDOLiteratureCrawler.transform, line 202
raw_html (read)	payload.get('raw_html')	GDOLiteratureCrawler.transform, line 200
citation_text (read)	payload.get('citation_text')	GDOLiteratureCrawler.transform, line 200
task_id (read)	kwargs.get('task_id')	GDOLiteratureCrawler.crawl, line 85

8. Research use, citation and rights

No general licence for redistributing the complete bibliography was established.[27,28]

S011. GdO Orgelbauer Links

Catalogue key: gdo_orgelbauer_links | Role: registered source | Documentary unit: Directory link to a builder, supplier or association

1. Identity, scope and evidence

Source account. The directory includes both individual firms and organ-building associations. Country headings organise the links. A firm's name and destination website form useful identifying evidence, while association links describe a different type of organisation.[31]

Attribution recorded in the catalogue: Gesellschaft der Orgelfreunde. Scope: International organ-building directory grouped by country.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The crawler treats an outbound directory link and its heading path as the acquisition unit. Link text, directory URL and target website are preserved separately.[32,33]

Configured source address: https://www.gdo.de/recherchen/links/orgelbauer. This is a registry setting, not a new availability test.

Configured primary identifier: institution_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

Normalization constructs an institution-directory entry, retains display label, country, locality hint and heading position, and supplies a derived identifier when necessary. The linked institution site is recorded as a candidate for later enrichment. That intended later work is not already part of the directory evidence.[32,33]

5. Dates, relationships and historical states

A directory appearance does not establish current trading status, legal continuity or responsibility for a historical instrument. Those questions require the destination's own records and dated construction evidence.[31]

6. Preservation, loss and uncertainty

Directory grouping does not prove an institution's legal identity, foundation date or present operation. A website URL and an organisation are different entities. Default organ-builder classification follows the directory context and should not silently become an independently verified occupation.[32,33]

7. Implementation evidence and rule register

Registered class: GDOOrgelbauerLinksCrawler in aggregator/gdo_orgelbauer_links_crawler.py; inheritance: GDOOrgelbauerLinksCrawler > BaseCrawler.[32,33]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GDOOrgelbauerLinksCrawler.crawl (lines 113-132).
- GDOOrgelbauerLinksCrawler.transform (lines 187-211).
- GDOOrgelbauerLinksCrawler.parse_directory_html (lines 139-185).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
country	country	GDOOrgelbauerLinksCrawler.parse_directory_html, line 170
raw_data_json (read)	payload.get('raw_data_json')	GDOOrgelbauerLinksCrawler.transform, line 189
heading_path (read)	payload.get('heading_path')	GDOOrgelbauerLinksCrawler.transform, line 198
crawl_run_id (read)	payload.get('crawl_run_id')	GDOOrgelbauerLinksCrawler.transform, line 209
transform_run_id (read)	payload.get('transform_run_id')	GDOOrgelbauerLinksCrawler.transform, line 210
raw_html (read)	payload.get('raw_html')	GDOOrgelbauerLinksCrawler.transform, line 208

8. Research use, citation and rights

The directory carries a society copyright notice; no general open licence was established.[31]

S012. GrandOrgue GitHub Wiki: Sample Sets

Catalogue key: grandorgue_github_wiki | Role: registered source | Documentary unit: Sample-set or organ-definition-file directory entry

1. Identity, scope and evidence

Source account. The wiki separates free or demonstration sample sets from organ definition files for commercial recordings. It names external hosts, including Lars Palo, MPS Orgelseite and Virtual Pipe Organs, and tells readers to consult individual licences.[34]

Attribution recorded in the catalogue: GrandOrgue project contributors. Scope: GrandOrgue resources across multiple producers.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects wiki heading hierarchy and list/paragraph nodes inside the wiki or Markdown body. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://github.com/GrandOrgue/grandorgue/wiki/Sample-Sets. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "registry" (aggregator/grandorgue_library_crawler.py, line 515).

4. Extraction and normalization

The implemented unit is heading context and individual list entries. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Several Aggregator keys describe resources also listed here. These are multiple discovery paths, not independent sample-set populations. A wiki revision is preferable to an undated link when recording changes in availability.[34]

6. Preservation, loss and uncertainty

Nested lists and headings can describe producers, licences or requirements rather than independent sample sets. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: GrandOrgueGitHubWikiCrawler in aggregator/grandorgue_library_crawler.py; inheritance: GrandOrgueGitHubWikiCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- GrandOrgueGitHubWikiCrawler.parse_html (lines 517-573).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'odf' if is_odf else 'registry_entry'	GrandOrgueGitHubWikiCrawler.parse_html, line 557
requires_sample_set	is_odf	GrandOrgueGitHubWikiCrawler.parse_html, line 558
is_odf_only	is_odf	GrandOrgueGitHubWikiCrawler.parse_html, line 559

8. Research use, citation and rights

Individual sample-set and ODF licences apply; the software project's licence is not evidence of the audio's licence.[34]

S013. INdO

Catalogue key: indo | Role: registered source | Documentary unit: Organ record with history and technical description

1. Identity, scope and evidence

Source account. Orgue en France develops and maintains the collaborative national inventory, supported by the French Ministry of Culture since 2017. The site opened to public contributions in 2021 and describes its purpose as extending earlier departmental and regional inventories.[35,36]

Attribution recorded in the catalogue: Orgue en France. Scope: France, including overseas departments and territories in the search interface.

2. Native record structure and contents

Selected document; 2026-09-11. Structured organ profile with linked authority records and contribution history[36]

- Building, town, department and position within the building are populated.
- The organ profile links to a location map.
- Builder and restorer entries have roles and dates.
- Construction, restoration and modification activities are listed.
- Eight named stops with pitches in the selected manual division.
- Manual compass and note count are supplied.
- Key and stop actions and windchest description are populated.
- Temperament and pitch include a temperature-qualified measurement.
- The bellows/reservoir arrangement is described.
- Case and console have separate descriptions.
- Playability and ownership are stated, with heritage-protection identifiers.
- The record contains linked instrument images.
- Bibliographic sources and a linked documentary extract are identified.
- National heritage identifiers and builder-record links are present.
- Update date, named contributor and contribution history are displayed.

Repeated structures: Sound divisions contain stops and compasses. Builder activities repeat with roles and dates.

3. Discovery and acquisition

The adapter acquires inventory detail pages and uses the cached town/building information as a fallback when the rendered heading does not supply it.[37]

Configured source address: https://inventaire-des-orgues.fr. This is a registry setting, not a new availability test.

Configured primary identifier: indo_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

Composition cards are grouped by keyboard or pedal; their list items become source stop strings. Card headings determine the group names and a pedal-text test determines the branch. Location is read from the page heading and locality links. Timeline/history and media information are handled separately from the composition cards. Later offline attribution helpers operate on preserved evidence and are not evidence that the original crawler performed actor reconciliation.[37]

5. Dates, relationships and historical states

Construction and restoration dates are distinct from the page revision and contributor history.[36]

Existing and disappeared instruments are separate categories. Construction, rebuilding and current condition need their own dates and attributions. An inventory entry's completion state is distinct from the physical condition of the organ.[35,36]

6. Preservation, loss and uncertainty

A stop string can contain pitch and annotations without having those parts independently resolved. Keyboard grouping and the distinction between an instrument and successive works remain source-specific. The France value is supplied by the adapter. Neither that default nor a successful transform measures the national inventory's completeness.[37]

Limit of the selected-document observation: Visible headings do not guarantee values. The examined barrel-organ record has no ordinary pedal division.

7. Implementation evidence and rule register

Registered class: IndoCrawler in aggregator/indo_crawler.py; inheritance: IndoCrawler > BaseCrawler.[37]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- IndoCrawler.crawl (lines 47-193).
- IndoCrawler.transform (lines 236-393).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
country	country	IndoCrawler.transform, line 273
raw_str	stop_name	IndoCrawler.transform, line 308
html (read)	raw_data.get('html', '')	IndoCrawler.transform, line 240
town (read)	raw_data.get('town', '')	IndoCrawler.transform, line 262
builder (read)	raw_data.get('builder')	IndoCrawler.transform, line 369
building (read)	raw_data.get('building', '')	IndoCrawler.transform, line 259

8. Research use, citation and rights

The consulted editorial page did not establish a single licence covering records, photographs and attachments.[35,36]

S014. IOHIO

Catalogue key: iohio | Role: registered source | Documentary unit: Historical organ profile

1. Identity, scope and evidence

Source account. IOHIO's bilingual website presents its work on the historic organs of Oaxaca. Aggregator registers it as an organ source; the public site also contains institutional and programme material beyond the organ profiles.[38,39,40]

Attribution recorded in the catalogue: Instituto de Órganos Históricos de Oaxaca. Scope: Oaxaca, Mexico.

2. Native record structure and contents

Selected document; 2026-09-11. Bilingual instrument profile with labelled technical summary and historical essay[40]

- Church and region identify the organ.
- Original maker is unknown; restorers and restoration participants are named.
- Construction, relocation, rebuilding and restoration are described with dates.
- Separate bass and treble stop groups identify cancelled reed stops.
- Keyboard range, short octave and split point are specified.
- Reconstruction of keyboard action and windchest is described.
- Pitch and meantone temperament are stated.
- Bellows construction and wind pressure are stated.
- Case measurements, decoration and keyboard configuration are documented.
- Current use and surviving original components are described.
- Case, pipe-decoration and detail photographs are included.
- The narrative refers to dated historical inventories.
- Comparable instruments are linked.

Repeated structures: Specification is split into left-hand and right-hand stop groups.

3. Discovery and acquisition

The adapter reads acquired Squarespace-style content blocks, labelled metadata and detail-page media. The source account concerns a specialist regional organ inventory rather than a uniform global register.[41]

Configured source address: https://www.iohio.org.mx. This is a registry setting, not a new availability test.

Configured primary identifier: iohio_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

Bold labels and following text supply a metadata dictionary. Builder and construction-date labels feed attribution; pitch has a dedicated field, while case dimensions, keyboard and bellows are retained as notes. Stop blocks are recognised through headings such as bass, treble and left or right hand. The parser therefore recognises a functional split within a manual rather than only conventional named divisions.[41]

5. Dates, relationships and historical states

Original construction, later alterations and restoration are distinguished in prose.[40]

Its institutional home page supplies context, but only an individual profile can substantiate an instrument's specification or conservation state.[38,39,40]

6. Preservation, loss and uncertainty

A heading-based split should not be mistaken for separate keyboards. Narrative conditions such as cancelled stops, alterations or restoration state require interpretation with the source block. The town can be taken from the page title and the building from reference context; those are different evidential origins.[41]

Limit of the selected-document observation: Split registers and cancelled stops need their own attributes; a flat list loses their functional meaning.

7. Implementation evidence and rule register

Registered class: IohioCrawler in aggregator/iohio_crawler.py; inheritance: IohioCrawler > BaseCrawler.[41]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- IohioCrawler.crawl (lines 46-143).
- IohioCrawler.transform (lines 211-371).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
dates	[build_date]	IohioCrawler.transform, line 278
country	'Mexico'	IohioCrawler.transform, line 262
raw_str	metadata['pitch']	IohioCrawler.transform, line 288
html (read)	raw_data.get('html', '')	IohioCrawler.transform, line 212
town (read)	raw_data.get('town', '')	IohioCrawler.transform, line 255
notes (read)	data['specifications'].get('notes', [])	IohioCrawler.transform, line 290

8. Research use, citation and rights

No general licence for reproducing the collection was established from the consulted pages.[38,39,40]

S015. Jeux d'orgues Sample Set Downloads

Catalogue key: jeux_d_orgues_downloads | Role: registered source | Documentary unit: Sample-set release and format variant

1. Identity, scope and evidence

Source account. The project records historic French organs and distributes products in several virtual-instrument and sound-bank formats. Its site lists distinct numbered releases, including Stiehr-Mockers, Silbermann, La Chaise-Dieu and Souvigny.[42]

Attribution recorded in the catalogue: Jeux d'orgues. Scope: Historic French organs.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects links matching the producer's numbered sample-set naming pattern. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://www.jeuxdorgues.com/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_downloads" (aggregator/grandorgue_library_crawler.py, line 1214).

4. Extraction and normalization

The implemented unit is download/product entries. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Product number, organ identity and software format are separate fields. A format named on the project home page should not be assigned to every release without checking that release's page.[42]

6. Preservation, loss and uncertainty

Filename or product numbering identifies a publication unit rather than a physical historical configuration. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: JeuxDOrguesDownloadsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: JeuxDOrguesDownloadsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- JeuxDOrguesDownloadsCrawler.parse_html (lines 1216-1263).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'sample_set'	JeuxDOrguesDownloadsCrawler.parse_html, line 1243
requires_sample_set	False	JeuxDOrguesDownloadsCrawler.parse_html, line 1246
is_odf_only	False	JeuxDOrguesDownloadsCrawler.parse_html, line 1245
producer_label	"Jeux d'orgues"	JeuxDOrguesDownloadsCrawler.parse_html, line 1234

8. Research use, citation and rights

The home page asserts copyright. Price and download availability do not resolve file-specific licence terms.[42]

S016. Kirchenorgel-Forum

Catalogue key: kirchenorgelforum | Role: registered source | Documentary unit: Discussion thread and dated post

1. Identity, scope and evidence

Source account. The forum divides pipe organs from electronic, digital and virtual instruments and provides further sections on scores, interpretation, improvisation and events. The public index was readable without entering an account.[43]

Attribution recorded in the catalogue: Kirchenorgel-Forum. Scope: Pipe, electronic and virtual organs; playing, music and community discussion.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

Forum pages can arrive as an explicit page list, a serialised multi-page capture or one HTML page. Index context and thread-page evidence are handled separately.[44,45]

Configured source address: https://www.kirchenorgelforum.at. This is a registry setting, not a new availability test.

Configured primary identifier: thread_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The transform constructs one thread record containing posts and forum context. It fills missing thread metadata from the listing, retains page URLs and captured-page count, and can attach a separately hashed index document. Post, quote, reaction, author, link and media extraction are distinct routines.[44,45]

5. Dates, relationships and historical states

Cite a post and its date for a contributor's statement. A thread title alone does not substantiate a technical claim, and a public index does not show that every linked post or attachment is accessible. This package contains no participant prose.[43]

6. Preservation, loss and uncertainty

One captured thread page does not establish thread completeness. Quoted text and a post author have different attribution. Statements in posts remain attributed claims; the forum is not an instrument inventory or an independent confirmation of a specification.[44,45]

7. Implementation evidence and rule register

Registered class: KirchenorgelForumCrawler in aggregator/kirchenorgelforum_crawler.py; inheritance: KirchenorgelForumCrawler > BaseCrawler.[44,45]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- KirchenorgelForumCrawler.crawl (lines 156-265).
- KirchenorgelForumCrawler.transform (lines 340-394).
- KirchenorgelForumCrawler._extract_images (lines 801-836).
- KirchenorgelForumCrawler._extract_links (lines 783-799).
- KirchenorgelForumCrawler._extract_quotes (lines 838-858).

- KirchenorgelForumCrawler._extract_reactions (lines 861-878).
- KirchenorgelForumCrawler._media_images_from_posts (lines 907-911).
- KirchenorgelForumCrawler._parse_forum_thread_rows (lines 571-618).
- KirchenorgelForumCrawler._parse_post (lines 522-569).
- KirchenorgelForumCrawler._parse_user_meta (lines 768-781).

The JSON companion lists the remaining 2 source-specific routines and the other eligible inherited methods.

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
source_snapshot	{'page_urls': [page['url'] for page in pages if page.get('url')]}	KirchenorgelForumCrawler.parse_thread_pages, line 518
html (read)	raw_data.get('html')	KirchenorgelForumCrawler.transform, line 341
raw_html (read)	raw_data.get('raw_html')	KirchenorgelForumCrawler.transform, line 341
url (read)	raw_data.get('url')	KirchenorgelForumCrawler.transform, line 345
forum_id (read)	forum.get('forum_id')	KirchenorgelForumCrawler.transform, line 358
forum_title (read)	forum.get('forum_title')	KirchenorgelForumCrawler.transform, line 359

8. Research use, citation and rights

No republication permission for post text, attachments or participant data was established.[43]

S017. Lars Palo Free GrandOrgue Samplesets

Catalogue key: lars_palo_vpo | Role: registered source | Documentary unit: Sample-set package, variant and patch

1. Identity, scope and evidence

Source account. The download table distinguishes original and extended stop counts and gives package versions, file sizes and software requirements. Its licensing statement separates older sets under CC BY-SA 2.5 from the 2023 Norrfjärden set under CC BY-SA 4.0.[46]

Attribution recorded in the catalogue: Lars Palo. Scope: Primarily Swedish organs prepared for GrandOrgue.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects table rows, then article/download/list nodes as a fallback. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://familjenpalo.se/vpo/download/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_downloads" (aggregator/grandorgue_library_crawler.py, line 791).

4. Extraction and normalization

The implemented unit is row-based items plus a specialised detail-enrichment step. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Version and original/extended status are necessary comparison fields. Summing the stop counts of virtual extensions would misdescribe the recorded instrument. A patch belongs to the release it modifies.[46]

6. Preservation, loss and uncertainty

Original and extended versions require separate edition labels; a file link does not establish that both versions reproduce the same specification. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: LarsPaloVpoCrawler in aggregator/grandorgue_library_crawler.py; inheritance: LarsPaloVpoCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- LarsPaloVpoCrawler.parse_html (lines 793-870).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	LarsPaloVpoCrawler.parse_html, line 857
manual_count	manual_count	LarsPaloVpoCrawler.parse_html, line 855
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'sample_set'	LarsPaloVpoCrawler.parse_html, line 841
requires_sample_set	False	LarsPaloVpoCrawler.parse_html, line 850
is_odf_only	False	LarsPaloVpoCrawler.parse_html, line 849

8. Research use, citation and rights

The page reports Creative Commons licences with differing versions; the downloaded package's own licence remains the item-level reference.[46]

S018. Lavender Audio Hereford Hauptwerk Sample Sets

Catalogue key: lavender_audio_hereford | Role: registered source | Documentary unit: Hereford sample-set edition

1. Identity, scope and evidence

Source account. The page distinguishes an evaluation set, 23-stop and 46-stop editions, and the complete 67-stop instrument. It also identifies an XL extension and gives April 2009 and February 2010 as sampling dates.[47]

Attribution recorded in the catalogue: Lavender Audio. Scope: One physical organ represented by multiple virtual editions.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects page links selected within the Hereford product context. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://www.lavenderaudio.co.uk/organs/hereford/hwss.html. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_catalog" (aggregator/grandorgue_library_crawler.py, line 1707).

4. Extraction and normalization

The implemented unit is edition and file references. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

For example, the inspected product construction supplies manual_count: 4 and pedal_present: True as constants, while its stop count is obtained from a title match. These have different evidential origins.

5. Dates, relationships and historical states

These products form a particularly clear example of one organ yielding several catalogue records. Sampling date, selected ranks and extended keyboard compass must remain separate from the historical organ's specification.[47]

6. Preservation, loss and uncertainty

Reduced and extended editions are separate virtual configurations associated with one physical instrument; their stop totals should not be pooled. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: LavenderAudioHerefordCrawler in aggregator/grandorgue_library_crawler.py; inheritance: LavenderAudioHerefordCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- LavenderAudioHerefordCrawler.parse_html (lines 1709-1786).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stops	LavenderAudioHerefordCrawler.parse_html, line 1751
manual_count	4	LavenderAudioHerefordCrawler.parse_html, line 1749
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'evaluation_sample_set' if is_eval else 'sample_set'	LavenderAudioHerefordCrawler.parse_html, line 1741
requires_sample_set	False	LavenderAudioHerefordCrawler.parse_html, line 1743
is_odf_only	False	LavenderAudioHerefordCrawler.parse_html, line 1742

8. Research use, citation and rights

The page retains copyright in samples, images and definition files and specifies licensed software use. The evaluation edition is not an open-data release.[47]

S019. Milan Digital Audio Instruments

Catalogue key: milan_digital_audio_instruments | Role: registered source | Documentary unit: Commercial virtual-organ product

1. Identity, scope and evidence

Source account. The indexed official catalogue identifies products based on organs at St Eucaire, Salisbury, Bovenkerk, Notre Dame de Metz, Mt Carmel and St Peter and Paul. Direct requests to the site returned HTTP 403 in this review.[48,49]

Attribution recorded in the catalogue: Milan Digital Audio. Scope: Selected historical and modern instruments for Hauptwerk.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects article elements and their product links. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://milandigitalaudio.com/instruments/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_catalog" (aggregator/grandorgue_library_crawler.py, line 1472).

4. Extraction and normalization

The implemented unit is instrument product articles. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

The indexed listing supports identification of the catalogue, not current prices, download availability or release compatibility. Those fields remain unverified. Record the product separately from the physical instrument.[48,49]

6. Preservation, loss and uncertainty

A product page can outlast the release it describes; format, version and availability need their own dated evidence. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: MilanDigitalAudioInstrumentsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: MilanDigitalAudioInstrumentsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- MilanDigitalAudioInstrumentsCrawler.parse_html (lines 1474-1519).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'sample_set'	MilanDigitalAudioInstrumentsCrawler.parse_html, line 1500
requires_sample_set	False	MilanDigitalAudioInstrumentsCrawler.parse_html, line 1502
is_odf_only	False	MilanDigitalAudioInstrumentsCrawler.parse_html, line 1501
producer_label	'Milan Digital Audio'	MilanDigitalAudioInstrumentsCrawler.parse_html, line 1491

8. Research use, citation and rights

Product licence terms were not accessible in the direct check.[48,49]

S020. MPS Orgelseite Filebase

Catalogue key: mps_orgelseite_filebase | Role: registered source | Documentary unit: Filebase entry, including ODFs, composites and other resources

1. Identity, scope and evidence

Source account. The current filebase mixes definition files, composite organs, simulations, tools and musical material. Entries expose dates, revisions and download counters, with separate GrandOrgue and Hauptwerk categories.[50]

Attribution recorded in the catalogue: MPS Orgelseite contributors. Scope: Community resources for GrandOrgue and Hauptwerk.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects filebase links, file IDs, article/file-entry containers and entry-title links. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://mps-orgelseite.de/home/filebase/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "filebase" (aggregator/grandorgue_library_crawler.py, line 579).

4. Extraction and normalization

The implemented unit is file entries and their detail/download context. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

A filebase category is not a homogeneous sample-set collection. Classify each entry before joining it to organs or products. Public metadata and downloadable bytes have different access conditions; no attachments were retrieved here.[50]

6. Preservation, loss and uncertainty

A filebase item may be an update, definition or supporting file; item identity is not an organ identity. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: MpsOrgelseiteFilebaseCrawler in aggregator/grandorgue_library_crawler.py; inheritance: MpsOrgelseiteFilebaseCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- MpsOrgelseiteFilebaseCrawler.parse_html (lines 581-646).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	package_kind	MpsOrgelseiteFilebaseCrawler.parse_html, line 628
requires_sample_set	is_odf or is_patch	MpsOrgelseiteFilebaseCrawler.parse_html, line 631
is_odf_only	is_odf	MpsOrgelseiteFilebaseCrawler.parse_html, line 630

8. Research use, citation and rights

Permissions depend on the contribution and on any underlying samples it requires.[50]

S021. Musique et musiciens - Orgues

Catalogue key: musiqueorguequebec | Role: registered source | Documentary unit: Organ profile within a broader music website

1. Identity, scope and evidence

Source account. The configured root identifies a bilingual music website rather than an organ-only register. Aggregator's organ adapter selects profiles from this larger resource. The root-page response by itself gives no reliable organ-record count.[51,1,52]

Attribution recorded in the catalogue: Musique et musiciens / Music and musicians. Scope: Organ-related pages; wider musical content on the parent site.

2. Native record structure and contents

Selected document; 2026-09-11. Bilingual church-and-organ essay with comparative specification tables[52]

- Church, municipality and province identify the page.
- Original builder/opus and a later builder are named.
- Dated building and organ narratives are separate sections.
- Original and current specifications are compared by division.
- Manual and pedal ranges and note counts are provided.
- Electro-pneumatic action is stated.
- Couplers, adjustable combinations and expression controls are listed.
- The page includes an instrument photograph and enlarged-image link.
- A dedicated bibliography closes the page.
- Internal language/specification anchors and image links are present.

Repeated structures: Original and current stops appear side by side within each division.

3. Discovery and acquisition

Index discovery leads to bilingual or French organ articles. The adapter parses page text and stop tables with source-specific routines.[5]

Configured source address: http://www.musiqueorguequebec.ca. This is a registry setting, not a new availability test.

Configured primary identifier: moq_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Canada" (aggregator/static_html_crawler.py, line 2802).
- start_urls: ["http://www.musiqueorguequebec.ca/orgues/orgueque3.html", "http://www.musiqueorguequebec.ca/orgues/orguecan3.html"] (aggregator/static_html_crawler.py, line 2803).
- detail_url_patterns: ["/orgues/quebec/", "/orgues/canada/"] (aggregator/static_html_crawler.py, line 2807).
- exclude_url_patterns: [".jpg", ".gif", ".png", ".pdf", "repertoiredesorgues.qc.ca"] (aggregator/static_html_crawler.py, line 2808).

4. Extraction and normalization

Builder history is parsed from narrative patterns. Manual, stop and rank totals are sought through explicit textual expressions; stop count falls back to the number of extracted entries. Stop tables are preserved as raw specification text and division groups, while action, images and location have separate extraction paths.[5]

5. Dates, relationships and historical states

Building history and organ history coexist. Original and later specifications are explicitly differentiated.[52]

Catalogue references should point to the organ profile, not the general home page. The site's name should not be used to infer that every documented instrument is in Québec.[51,1,52]

6. Preservation, loss and uncertainty

Parallel original and later specifications need state-specific reading. A grouped stop dictionary does not automatically preserve the semantics of parallel historical columns. Declared totals and sums of parsed entries should be reported with their respective origins.[5]

Limit of the selected-document observation: Parallel original/current columns must remain aligned; church events must not become organ construction events.

7. Implementation evidence and rule register

Registered class: MusiqueOrgueQuebecCrawler in aggregator/static_html_crawler.py; inheritance: MusiqueOrgueQuebecCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- MusiqueOrgueQuebecCrawler.transform (lines 2849-2895).
- MusiqueOrgueQuebecCrawler._builder_history (lines 2907-2920).
- MusiqueOrgueQuebecCrawler._extract_stop_tables (lines 2930-2958).
- MusiqueOrgueQuebecCrawler._location_from_title (lines 2897-2905).
- MusiqueOrgueQuebecCrawler._parse_index_page (lines 2810-2844).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count or sum((len(stops) for stops in stops_by_division.values()))	MusiqueOrgueQuebecCrawler.transform, line 2863
manual_count	manual_count	MusiqueOrgueQuebecCrawler.transform, line 2862
rank_count	rank_count	MusiqueOrgueQuebecCrawler.transform, line 2864
stops_by_division	stops_by_division	MusiqueOrgueQuebecCrawler.transform, line 2868
raw_specification	raw_spec	MusiqueOrgueQuebecCrawler.transform, line 2866
country	self.default_country	MusiqueOrgueQuebecCrawler.transform, line 2875

8. Research use, citation and rights

No general licence for the selected organ profiles was established.[51,1,52]

S022. musiXplora

Catalogue key: musixplora | Role: registered source | Documentary unit: Typed entity and relation

1. Identity, scope and evidence

Source account. The public interface separates seven lexical domains. Its person search exposes several external authority identifiers, while the project description presents musiXplora as an expanding organological and musicological reference resource.[53,1]

Attribution recorded in the catalogue: musiXplora / Universität Leipzig. Scope: Persons, institutions, objects or works, places, events, concepts and titles.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The crawler facade loads local dump tables through MusiXploraGraphBuilder, builds source-native entities and caches references. It does not acquire the public website in this path. Dump provenance and the public portal description therefore represent different evidence.[54,55,56,57]

Configured source address: https://musixplora.de. This is a registry setting, not a new availability test.

Configured primary identifier: mxp_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The normalizer receives an entity and an optional manifest. Graph construction, entity normalization and an optional pipe-organ projection are separate layers. Relations, entity family/type and source IDs belong to the dump model; organ relevance is a selection assessment rather than an entity identity.[54,55,56,57]

5. Dates, relationships and historical states

Entity and relation identifiers can support comparison across collections. Observations of the public interface do not establish the coverage or date of a separately obtained data extract.[53,1]

6. Preservation, loss and uncertainty

A public portal page cannot independently verify every field in a local dump. Entity, relationship and organ-projection counts have different denominators. The organ projection must not erase non-organ entities or attribute all graph relations to physical instruments.[54,55,56,57]

7. Implementation evidence and rule register

Registered class: MusiXploraCrawler in aggregator/musixplora/crawler.py; inheritance: MusiXploraCrawler > BaseCrawler.[54,55,56,57]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- MusiXploraCrawler.crawl (lines 51-111).
- MusiXploraCrawler.transform (lines 113-116).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
entity (read)	raw_data.get('entity')	MusiXploraCrawler.transform, line 114
manifest (read)	raw_data.get('manifest')	MusiXploraCrawler.transform, line 115
task_id (read)	kwargs.get('task_id')	MusiXploraCrawler.crawl, line 54
limit (read)	kwargs.get('limit')	MusiXploraCrawler.crawl, line 55
index_limit (read)	self.config.get('index_limit')	MusiXploraCrawler.crawl, line 55
source_dir (read)	kwargs.get('source_dir')	MusiXploraCrawler.crawl, line 53

8. Research use, citation and rights

Open access is stated by the project; a licence for redistributing the local dump was not established by the public pages.[53,1]

S023. Nomine

Catalogue key: nomine | Role: registered source | Documentary unit: Regional organ portrait

1. Identity, scope and evidence

Source account. NOMINE organises portraits by East Frisia, Oldenburg, Stade and Lüneburg and links the instruments to recordings and video portraits. Its introduction describes a regional heritage of 300 historic organs.[58,59]

Attribution recorded in the catalogue: Norddeutsche Orgelmusikkultur in Niedersachsen und Europa e.V. (NOMINE). Scope: Historical organs between the Ems and Elbe.

2. Native record structure and contents

Selected document; 2026-09-11. Regional organ portrait with explanatory prose, stop table and dated work history[59]

- Church, place and navigation address are supplied.
- Builders and their interventions are named.
- A chronological work history accompanies the present specification.
- Stops retain division, pitch and mixture-rank information.
- Manual/pedal compass and division labels are given.
- Historical windchest and action changes are discussed.
- Surviving case and coupling arrangements are described.
- Retention of earlier case and registers is explicit.
- Dated bibliography and literature citations are supplied.
- A documentary stand date is attached to the account.

Repeated structures: Division columns, technical rows and a separate construction/restoration chronology.

3. Discovery and acquisition

The acquired organ page is interpreted through its title, disposition heading, construction/restoration-history heading and uploaded images.[60]

Configured source address: https://nomine.net. This is a registry setting, not a new availability test.

Configured primary identifier: nomine_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The title is split into town and building. The first table reached from the disposition heading is passed to a specialised parser. History-table rows provide dates and participant wording. The image filter selects upload paths while excluding certain logo, icon and thumbnail patterns. Manual, pedal, stop and other specification groups are assembled separately.[60]

5. Dates, relationships and historical states

Current instrument, surviving earlier case and retained ranks are explicitly distinguished.[59]

The stated regional total is not a verified count of downloadable records. The combination of portraits and recordings supports contextual study, while a representative inventory would require a separate account of selection and omissions.[58,59]

6. Preservation, loss and uncertainty

History cells are retained as participant wording rather than independently resolved organisations or people. A next-table search can cross a section boundary if the expected table is absent. The disposition and history need their own state interpretation; the title split and default country are implementation choices.[60]

Limit of the selected-document observation: Count surviving components separately from the date of the present instrument.

7. Implementation evidence and rule register

Registered class: NomineCrawler in aggregator/nomine_crawler.py; inheritance: NomineCrawler > BaseCrawler.[60]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- NomineCrawler.crawl (lines 48-165).
- NomineCrawler.transform (lines 249-338).
- NomineCrawler._parse_disposition_table (lines 340-504).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
dates	[year]	NomineCrawler.transform, line 313
country	'Germany'	NomineCrawler.transform, line 279
raw_str	f'{year}: {builder_text}'	NomineCrawler.transform, line 314
html (read)	raw_data.get('html', '')	NomineCrawler.transform, line 250
src (read)	img.get('src')	NomineCrawler.transform, line 321
nomine_id (read)	raw_data.get('nomine_id')	NomineCrawler.transform, line 334

8. Research use, citation and rights

The site marks its content as protected; no open licence for text, photographs or recordings was established.[58,59]

S024. NPOR

Catalogue key: npor | Role: registered source | Documentary unit: Source survey linked to an organ

1. Identity, scope and evidence

Source account. BIOS provides the register, whose content is managed under an agreement with the Royal College of Organists. The current interface advertises more than 37,000 organs and links to the Historic Organ Sound Archive, British Organ Archive and Directory of British Organ Builders.[61,62]

Attribution recorded in the catalogue: British Institute of Organ Studies (BIOS). Scope: National Pipe Organ Register and linked British organ resources.

2. Native record structure and contents

Selected document; 2026-09-11. Survey record with separate history, builder and archive links[62]

- Church address, postcode and building type are supplied.
- National Grid reference links to a neighbourhood search.
- Named builder linked to a builder record.
- Relocation/rebuilding account links to earlier and related surveys.
- Department tables provide stop names, pitches and explanatory notes.
- Department-specific key compasses and note counts are populated.
- Key and stop action are stated separately for departments.
- Equal temperament is specified.
- Electric blowing is recorded.
- Separate case and console descriptions include control details.
- The survey marks the organ playable.
- Dated, credited case and console photographs are available.
- A British Organ Archive reference and source codes are supplied.
- Related survey, builder and history identifiers are linked.
- Survey date, input date, source code and dated updates are separate.

Repeated structures: Departments contain action, compass and stop rows. Survey updates and photographs are separate repeatable sections.

3. Discovery and acquisition

Discovery supports historic-county searches and explicit survey identifiers. Search results contain survey references; the detail request is keyed by the survey identifier. Search-reference context is retained alongside detail data, and disagreement in location fields is recorded.[63]

Configured source address: https://npor.org.uk. This is a registry setting, not a new availability test.

Configured primary identifier: `npor_id`. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The transform separates building and place identifiers from the survey identifier. Builders are iterated as source work assertions, with their ordinal, wording, date and attribution retained. Linked history remains documentary context; editorial updates remain record updates. Archive citations have their own extraction routine. Specification, photographic and bibliographic arrays are handled separately. A fixed normalization epoch is used by the inspected release policy, while acquisition dates belong to the surrounding evidence.[63]

5. Dates, relationships and historical states

Survey date, last update, building history, work dates and photograph dates are distinct.[62]

A survey identifier identifies a documentary record. Repeated surveys or revised descriptions must not be counted automatically as different physical organs. The linked archives are additional resources, not separate registered Aggregator keys in this inventory.[61,62]

6. Preservation, loss and uncertainty

Survey date, editorial update date and organ-work date have different meanings. A survey is not automatically one instrument over its entire lifetime. The code can establish the handling of builder assertions and source references, but cannot validate the historical accuracy of their wording. Detail/search conflicts should be examined rather than resolved by silently accepting the first value.[63]

Limit of the selected-document observation: Survey identifiers distinguish accounts. Related surveys may describe earlier organs or previous locations.

7. Implementation evidence and rule register

Registered class: NPORCrawler in aggregator/npor_crawler.py; inheritance: NPORCrawler > BaseCrawler.[63]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- NPORCrawler.crawl (lines 87-109).
- NPORCrawler.transform (lines 537-605).
- NPORCrawler._extract_specifications (lines 771-853).
- NPORCrawler._extract_events (lines 702-734).
- NPORCrawler._extract_literature (lines 896-938).
- NPORCrawler._event_type (lines 686-700).
- NPORCrawler._extract_descriptions (lines 736-769).
- NPORCrawler._extract_involved (lines 641-655).
- NPORCrawler._extract_locations (lines 624-639).
- NPORCrawler._extract_main_builder (lines 350-357).

The JSON companion lists the remaining 2 source-specific routines and the other eligible inherited methods.

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	str(stop_count)	NPORCrawler._extract_specifications, line 847
manual_count	str(manual_count)	NPORCrawler._extract_specifications, line 846
dates	[date] if date else []	NPORCrawler._extract_events, line 717
record_type	survey.get('record_type', '')	NPORCrawler.transform, line 561
country	'United Kingdom'	NPORCrawler._extract_locations, line 637
raw_str	'; '.join(citation_parts)	NPORCrawler._extract_literature, line 920

8. Research use, citation and rights

Copyright is displayed. No general open licence for surveys, archive images or sound recordings was established.[61,62]

S025. NYC AGO

Catalogue key: nycago | Role: registered source | Documentary unit: Building or venue page containing organ histories

1. Identity, scope and evidence

Source account. The project's official index describes an ongoing record of organs installed in the five boroughs. It distinguishes non-extant entries and indicates whether a specification is available. Direct requests were refused, so the content check used an indexed official page.[64,65,66]

Attribution recorded in the catalogue: New York City Chapter, American Guild of Organists. Scope: Present and former organs in all five New York City boroughs.

2. Native record structure and contents

Selected document; 2026-09-11. [66]

Access finding. The selected detail URL returned HTTP 403. Its native content was not examined. No native detail structure is inferred from this failed check.

3. Discovery and acquisition

The crawler discovers entries from the NYC AGO directory and retains listing title, category and borough with the detail HTML.[67]

Configured source address: https://nycago.org. This is a registry setting, not a new availability test.

Configured primary identifier: nycago_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The inspected transform obtains the venue from a marked span or bold element, scans marked builder spans and the following text for a four-digit year, and collects image references. It creates building/location, builder-history and source-credit structures. Its specifications object explicitly states that further parsing logic is required; it does not implement a detailed stop-table projection in this entry point.[67]

5. Dates, relationships and historical states

One building page may describe successive instruments. Borough, venue and installation history should therefore be retained when reconciling identifiers. The indexed evidence is older than this package's retrieval date.[64,65,66]

6. Preservation, loss and uncertainty

The archive can contain detailed stoplists even when this transform does not structure them. Years near builder names and venue-level pages do not settle the identity of successive organs. Image collection at this level is broad and does not by itself distinguish instrument photographs from every other image.[67]

7. Implementation evidence and rule register

Registered class: NycagoCrawler in aggregator/nycago_crawler.py; inheritance: NycagoCrawler > BaseCrawler.[67]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- NycagoCrawler.crawl (lines 258-293).
- NycagoCrawler.transform (lines 334-399).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
country	'United States'	NycagoCrawler.transform, line 379

Key / role	Inspected expression	Code locator
src (read)	img.get('src')	NycagoCrawler.transform, line 359
borough (read)	raw_data.get('borough')	NycagoCrawler.transform, line 378
title (read)	raw_data.get('title', '')	NycagoCrawler.transform, line 386
category (read)	raw_data.get('category', '')	NycagoCrawler.transform, line 387
link (read)	raw_data.get('link')	NycagoCrawler.transform, line 396

8. Research use, citation and rights

No collection-wide reuse permission was established from the consulted evidence.[64,65,66]

S026. OHS Pipe Organ Database

Catalogue key: ohs | Role: registered source | Documentary unit: Instrument record with related builders, places and documentary items

1. Identity, scope and evidence

Source account. The database exposes instrument, builder and location searching alongside stoplists, images and documents. Its home page separately reports totals for these entities and identifies the project as belonging to the Organ Historical Society.[68,69,70]

Attribution recorded in the catalogue: Organ Historical Society. Scope: Pipe Organ Database.

2. Native record structure and contents

Selected document; 2026-09-11. Relational instrument page with separately identified location, builder, stoplist and console records[70]

- Building, room, street address, city and location-record link.
- A map locates the installation.
- Named manufacturer linked to a builder record, with opus and year.
- Dated notes record later alterations and information updates.
- Selectable stoplist with named divisions, pitches and pipe quantities.
- Console record gives manual and pedal note counts.
- Key and stop actions are separate console attributes.
- Separate console record describes position, design and controls.
- Extant and playable are explicit instrument states.
- Three documents are counted, and the stoplist has an attributed documentary source.
- Location, builder and stoplist identifiers have separate routes.
- Original stoplist publication, receipt and later editorial dates are distinguished.

Repeated structures: Selectable stoplists and consoles; divisions contain stop rows. Documents, audio and notes are separate collections.

Not established in this document check: Image section and category filters are exposed; image payloads were not inspected.; The selected instrument explicitly reports zero audio items.; The selected instrument explicitly reports zero blower records..

3. Discovery and acquisition

An instrument request is supplemented by related endpoints and separately fetched stoplist resources. The crawler has a dedicated branch for uploaded text stoplists and a different branch for structured stoplist responses. Resource URLs and hashes accompany text parsing.[71,72,73]

Configured source address: https://pipeorgandatabase.org. This is a registry setting, not a new availability test.

Configured primary identifier: ohs_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

Instrument, location, builder, console, stoplist and media material enter different routines. The stoplist parser retains raw text, stops, couplers, accessories and unresolved lines. Console or division compass context may be attached to stop entries with a named provenance field. Current and original builder participation is delegated to a source-specific helper. The resulting instrument transfer is an assembly of several documentary objects, not a verbatim copy of one page.[71,72,73]

5. Dates, relationships and historical states

Instrument year, stoplist publication date, submission date and dated editorial notes are separate.[70]

The separate counters should remain separate populations. A stoplist may describe a particular date or configuration, and a builder attribution may concern an original instrument or later work. Neither should be collapsed into an undated present-day description.[68,69,70]

6. Preservation, loss and uncertainty

Failure of a related endpoint can be represented by an empty array; failure to fetch a stoplist is logged and the loop continues. Absence in the assembled transfer therefore does not prove absence upstream. A compass-based note count is contextual assignment, not a measured pipe count. Stoplist-resource dates and instrument chronology need to remain separately attributable.[71,72,73]

Limit of the selected-document observation: A stoplist or console is a child record with its own evidence. Zero audio items in this example is not a site-wide absence.

7. Implementation evidence and rule register

Registered class: OhsCrawler in aggregator/ohs_crawler.py; inheritance: OhsCrawler > BaseCrawler.[71,72,73]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OhsCrawler.crawl (lines 86-164).
- OhsCrawler.transform (lines 416-475).
- OhsCrawler._fetch_full_record (lines 361-401).
- OhsCrawler._parse_stoplists (lines 682-733).
- OhsCrawler._build_events (lines 528-546).
- OhsCrawler._build_literature (lines 781-788).
- OhsCrawler._build_location (lines 477-500).
- OhsCrawler._build_specifications (lines 593-680).
- OhsCrawler._coupler_label (lines 936-938).
- OhsCrawler._decode_stoplist_text (lines 404-410).

The JSON companion lists the remaining 3 source-specific routines and the other eligible inherited methods.

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count or None	OhsCrawler._build_specifications, line 656
manual_count	manual_count	OhsCrawler._build_specifications, line 655
pipe_count	self._first_int(item.get('pipes'))	OhsCrawler._build_specifications, line 660
rank_count	self._first_int(item.get('ranks'))	OhsCrawler._build_specifications, line 659
stops_by_division	stops_by_division	OhsCrawler._build_specifications, line 669
manual_compass	compass.get('manual')	OhsCrawler._build_specifications, line 661

8. Research use, citation and rights

The home page states that its information, images and documents may be used for educational or other purposes. Individual credits and rights still need to accompany reused items.[68,69,70]

S027. OHTA Organs of Australia

Catalogue key: ohta | Role: registered source | Documentary unit: Gazetteer entry and associated documentation

1. Identity, scope and evidence

Source account. OHTA combines gazetteers with conservation guidance, builder documentation and a separate John Stiller documentation project dated 1978-1986. It also lists relocated and redundant organs.[74,75,76]

Attribution recorded in the catalogue: Organ Historical Trust of Australia. Scope: Australian pipe-organ heritage.

2. Native record structure and contents

Selected document; 2026-09-11. Gazetteer article with footnoted historical essay and specification[76]

- Church and locality are identified in the article.
- Earlier and present builders and an opus are named.
- Construction, relocation and replacement are dated.
- Present instrument has a division-based stoplist with pitches.
- Manual/pedal note counts and division labels are stated.
- Mechanical key and stop action is stated.
- Equal temperament is stated.
- Case material, key coverings and pedalboard design are described.
- Removal and later installation of the predecessor are documented.
- Dated and attributed church, organ and console photographs.
- Numbered notes cite publications and on-site observations.
- Links connect the relocated instrument and related venues.
- Named author, observation dates and page-update date are supplied.

Repeated structures: Earlier and present organs occupy separate sections; stops are grouped by manual and pedal.

3. Discovery and acquisition

Index links lead to acquired organ articles; the adapter applies the static-page family and source-specific location/title handling.[5]

Configured source address: https://ohta.org.au. This is a registry setting, not a new availability test.

Configured primary identifier: ohta_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Australia" (aggregator/static_html_crawler.py, line 872).
- start_urls: ["https://ohta.org.au/organs/organs/QueenslandOrganHistory.html", "https://ohta.org.au/gazetteers/gazetteer-victorian-pipe-organs/", "https://ohta.org.au/gazetteers/gazetteer-queensland-pipe-organs/", "https://ohta.org.au/gazetteers/gazetteer-tasmanian-pipe-organs/", "https://ohta.org.au/gazetteers/gazetteer-nsw-pipe-organs/"] (aggregator/static_html_crawler.py, line 873).
- detail_url_patterns: ["/organs/"] (aggregator/static_html_crawler.py, line 880).
- exclude_url_patterns: ["queenslandorganhistory.html#", "organs-of-australia", "historical-overviews", "pipe-organ-documentation-project", ".pdf"] (aggregator/static_html_crawler.py, line 881).

4. Extraction and normalization

The common transform reads page text, metadata, specification candidates and images. The OHTA override checks whether a reference year is actually accompanied by the word opus and removes matching event/participant projections with an explicit warning. Its location rule interprets a comma-separated title as building followed by town.[5]

5. Dates, relationships and historical states

The selected article separates a removed predecessor from the present instrument and records its later destination.[76]

The historical survey and later gazetteer entries are distinct documentary layers. Relocation is especially important when matching an organ by its present address. Page authorship and the date of the underlying inspection should be retained where supplied.[74,75,76]

6. Preservation, loss and uncertainty

An article may document predecessors and successors. The opus safeguard addresses one known ambiguity, not every misleading number or historical date. The generic first-year fallback and construction projection still need checking against the article when no more explicit metadata is available.[5]

Limit of the selected-document observation: One venue article can describe multiple organs. Count instrument states and relocations separately from page count.

7. Implementation evidence and rule register

Registered class: OhtaCrawler in aggregator/static_html_crawler.py; inheritance: OhtaCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- OhtaCrawler.transform (lines 895-903).
- OhtaCrawler._location_parts (lines 905-909).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
year (read)	raw_data.get('year')	OhtaCrawler.transform, line 897
events (read)	record.get('events', [])	OhtaCrawler.transform, line 900
involved (read)	record.get('involved', [])	OhtaCrawler.transform, line 901
html (read)	raw_data.get('html')	OhtaCrawler.transform, line 898
raw_html (read)	raw_data.get('raw_html')	OhtaCrawler.transform, line 898
dates (read)	event.get('dates', [])	OhtaCrawler.transform, line 900

8. Research use, citation and rights

The published policy limits use to personal, educational and non-commercial purposes and requires permission for reproduction.[74,75,76]

S028. Organ Index

Catalogue key: organindex | Role: registered source | Documentary unit: MediaWiki article and revision

1. Identity, scope and evidence

Source account. Aggregator includes this organ source in its selection frame. The public entry point returned HTTP 403 during the review; its current articles, statistics and licence were not verified.[77,1,78]

Attribution recorded in the catalogue: Organ Index contributors. Scope: Organ descriptions in a collaborative wiki.

2. Native record structure and contents

Selected document; 2026-09-11. [78]

Access finding. The configured entry URL returned HTTP 403. No article was inspected. No native detail structure is inferred from this failed check.

3. Discovery and acquisition

The acquisition route follows wiki categories, category continuations and linked detail pages. The code also provides import routes for already captured category and detail HTML. Page title and category membership contribute to discovery; they are not organ identities across time.[79]

Configured source address: https://organindex.de. This is a registry setting, not a new availability test.

Configured primary identifier: organindex_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The detail transform reads a metadata table through a label map, constructs location and builder information, retains history text, and parses division columns below disposition headings. Images, literature, external links and categories have separate extractors. The disposition routine selects the next table after a matching heading and exits after the first successful parse.[79]

5. Dates, relationships and historical states

A revision identifier is preferable to an unversioned article URL. Article titles can change, and shared text or images may originate elsewhere. The failed request is an access observation, not evidence that the wiki has ceased operation.[77,1,78]

6. Preservation, loss and uncertainty

That first-success rule can omit additional specifications from the structured stoplist even when the page contains them. The table routine expects matching header and value cells in its first two direct rows. Other layouts require separate inspection. A wiki page can be richer than the fields populated by this parser; revision history is not established by a title or one captured page.[79]

7. Implementation evidence and rule register

Registered class: OrganIndexCrawler in aggregator/organindex_crawler.py; inheritance: OrganIndexCrawler > BaseCrawler.[79]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrganIndexCrawler.crawl (lines 149-226).
- OrganIndexCrawler.transform (lines 518-563).
- OrganIndexCrawler._extract_specifications (lines 659-691).
- OrganIndexCrawler._extract_events (lines 645-651).

- OrganIndexCrawler._extract_literature (lines 766-781).
- OrganIndexCrawler._parse_disposition_table (lines 710-728).
- OrganIndexCrawler._build_location (lines 601-619).
- OrganIndexCrawler._clean_stop_text (lines 851-857).
- OrganIndexCrawler._extract_categories (lines 783-789).
- OrganIndexCrawler._extract_descriptions (lines 653-657).

The JSON companion lists the remaining 8 source-specific routines and the other eligible inherited methods.

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	OrganIndexCrawler._extract_specifications, line 669
manual_count	manual_count	OrganIndexCrawler._extract_specifications, line 668
stops_by_division	divisions	OrganIndexCrawler._extract_specifications, line 672
dates	[metadata.get('year')] if metadata.get('year') else []	OrganIndexCrawler._extract_involved, line 640
country	'Germany'	OrganIndexCrawler._parse_category_page, line 298
raw_str	address or title	OrganIndexCrawler._build_location, line 618

8. Research use, citation and rights

Current wiki and image licence terms remain unverified in this review.[77,1,78]

S029. Orgbase

Catalogue key: orgbase | Role: registered source | Documentary unit: Organ description and specification

1. Identity, scope and evidence

Source account. Orgbase describes instruments worldwide and provides searches by builder, place, building, construction year and text. Its information page dates the database's beginning to October 1999. DBIS, the library database directory, reports more than 75,000 records; the consulted entry supplies no date for that quantity. The home page explicitly notes that not all descriptive information is available in every language.[80,81,82,83]

Attribution recorded in the catalogue: Orgbase / Organ Database. Scope: International; multilingual interface.

2. Native record structure and contents

Selected document; 2026-09-11. Searchable database with parameterised description records and technical sections[83]

- Country, region, city and named concert venue are populated.
- Named builder and construction year.
- Builder year and opening account provide dated instrument history.
- Stops are listed as comma-separated entries within divisions.
- Division labels and manual/pedal compasses are supplied.
- Separate key action, stop action and windchest fields.
- Pitch and temperament are separately populated.
- Wind pressure is stated with a unit.
- Couplers and a combination-system accessory are specified.
- Instrument photograph with named builder credit.
- A bibliography gives journal titles, issue details and authors.
- The venue website and source description identifier are present.
- The image supplier is credited; bibliography attribution is item-specific.

Repeated structures: Division stop counts precede prose-form division specifications, couplers and bibliography.

3. Discovery and acquisition

Country-list discovery and detail-page transformation are separate operations. A list row supplies an identifier and location, builder and manual-summary context; the HTML detail supplies the documentary sections. The offline transformer accepts already acquired evidence, so acquisition and interpretation can be studied separately.[84]

Configured source address: http://orgbase.nl. This is a registry setting, not a new availability test.

Configured primary identifier: orgbase_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The transform partitions the page into identification and location, activity tables, technical information, alternative specifications, footer references and images. Location histories preserve source-row positions. The alternative-specification branch collects labelled groups and their divisions; the compound-stop branch is distinct from ordinary stop parsing. Footer literature is merged with table literature rather than replacing it. A missing detail location can be filled from the listing context. These fallbacks and merges are part of the interpretation and must remain distinguishable from literal source statements. Technical labels are mapped individually: total stops, key action, stop action, windchests, temperament, pitch, wind pressure and division totals. Unrecognised technical rows are retained. At stop level the parser separates date strings, parenthetical notes, mixture-rank wording and pipe lengths from a cleaned name. The first detected pipe length supplies a MIDI conversion, and the resulting number is used to calculate a sounding transposition.[84]

5. Dates, relationships and historical states

Builder date and documentary literature dates are distinct. Language versions may differ in completeness.[83]

The DBIS figure establishes a reported scale, not a verified current total or a count of distinct surviving instruments. Language-specific versions may be incomplete relative to one another. Searches and comparisons should preserve the original language and source identifier. The information page's first-person account supplies a history but no named author in the retrieved text.[80,81,82,83]

6. Preservation, loss and uncertainty

A source record can therefore carry several places and specifications. Do not assign every stop to one present-day configuration. The location parser skips rows without a recognised year expression; its retention rule is narrower than all rows containing meaningful prose. A parsed division, stop or activity is not an independently identified physical entity. The presence of alternative-specification code establishes implemented handling, not that every historical layout is correctly parsed. The B/D branch supplies fixed bass and discant MIDI intervals, [36, 59] and [60, 96]. Those intervals are parser defaults, not measurements of that organ's split point or compass. A No stops present input is converted to an unavailable/unknown placeholder. The cleaned stop name, derived pitch representation and preserved wording consequently have different evidential status.[84]

Limit of the selected-document observation: Comma-separated division text includes pitches, mixture ranks, borrowing and compass information. It is not a uniform row-per-stop table.

7. Implementation evidence and rule register

Registered class: OrgbaseCrawler in aggregator/orgbase_crawler.py; inheritance: OrgbaseCrawler > BaseCrawler.[84]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrgbaseCrawler.crawl (lines 209-362).
- OrgbaseCrawler.transform (lines 619-726).

The JSON companion lists the remaining 30 source-specific routines and the other eligible inherited methods.

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
Total number of stops	'stop_count'	OrgbaseCrawler._structure_technical_data, line 964
Key action	'key_action'	OrgbaseCrawler._structure_technical_data, line 965
mixture	clean_string_by_ls_elements(mixture_detected, mixture_identifiers)	OrgbaseCrawler._process_stop_specs, line 1938
range_mapped	convert_pipelength_to_midi(length=primary_length, fractal=primary_fractal)	OrgbaseCrawler._process_stop_specs, line 1986
transposition_sounding	int(stop_spec_data['range_mapped'][0] - 36)	OrgbaseCrawler._process_stop_specs, line 1997
range	[36, 59]	OrgbaseCrawler._process_stop_specs, line 2003

8. Research use, citation and rights

No general open licence was established from the home and information pages.[80,81,82,83]

S030. Orgel-Archiv Schmidt

Catalogue key: orgelarchiv_schmidt | Role: registered source | Documentary unit: Organ profile

1. Identity, scope and evidence

Source account. The site's history traces the family archive to Horst Schmidt in 1972 and its continuation by Andreas Schmidt from 1995. Andreas Schmidt and Stefan Weyreder established the website. The account stresses personal inspection and the identification of reliable external sources.[85,86,87]

Attribution recorded in the catalogue: Andreas Schmidt; Stefan Weyreder. Scope: Germany, Switzerland, Austria and further country sections.

2. Native record structure and contents

Selected document; 2026-09-11. Contributed organ article with labelled location fields and distinct history sections[87]

- Church, denomination, place and administrative geography are labelled.
- Builders are named in the organ chronology.
- Dated construction, rebuilding and damage history.
- Stop groups explicitly include vacant positions.
- Division headings include key ranges.
- Historical action and windchest type are specified.
- Detailed controls and console fittings are described.
- A dated statement says the selected organ is unplayable.
- Photographic gallery and named contributor.
- Sources distinguish on-site inspection from published building history.
- A referenced heritage article is linked.
- Named author and inspection year are attached to the contribution.

Repeated structures: Manual/pedal stop groups, playing aids, building history and source credits.

3. Discovery and acquisition

The adapter has a WordPress-oriented discovery route and acquires the article body with its labels, history, disposition and source sections.[5,88]

Configured source address: https://orgel-verzeichnis.de. This is a registry setting, not a new availability test.

Configured primary identifier: orgelarchiv_schmidt_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Deutschland" (aggregator/static_html_crawler.py, line 3608).
- start_urls: ["https://orgel-verzeichnis.de/wp-json/wp/v2/posts"] (aggregator/static_html_crawler.py, line 3609).

4. Extraction and normalization

Section boundaries are explicitly defined. Disposition text is split into versions; the first version supplies the principal specification and the remaining versions are preserved as historical specifications. Declared and parsed stop counts are stored separately, and vacant positions are retained. History is passed to a source-specific participant helper rather than silently becoming resolved people or organisations.[5,88]

5. Dates, relationships and historical states

Organ history, church history, inspection year and condition date are separately stated.[87]

This stated editorial practice is relevant to source criticism, but it is not an independent accuracy audit. Instrument-level credits and inspection dates are still needed when using an entry as evidence.[85,86,87]

6. Preservation, loss and uncertainty

The first-version choice is an implementation convention that needs to match the article's ordering. Vacant positions are not sounding stops. Parsed totals can differ from declared historical totals, especially where several versions occur. Preserving raw history does not resolve ambiguous participant identities.[5,88]

Limit of the selected-document observation: The example distinguishes planned but unbuilt stops from installed stops. Vacancies must not be counted as sounding registers.

7. Implementation evidence and rule register

Registered class: OrgelArchivSchmidtCrawler in aggregator/static_html_crawler.py; inheritance: OrgelArchivSchmidtCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5,88]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrgelArchivSchmidtCrawler.crawl (lines 3617-3671).
- OrgelArchivSchmidtCrawler.transform (lines 3697-3760).
- OrgelArchivSchmidtCrawler._builder_history_from_text (lines 3804-3806).
- OrgelArchivSchmidtCrawler._extract_schmidt_title (lines 3762-3767).
- OrgelArchivSchmidtCrawler._is_schmidt_stop_line (lines 3825-3836).
- OrgelArchivSchmidtCrawler._manual_count_from_spec (lines 3864-3870).
- OrgelArchivSchmidtCrawler._stop_count_from_history (lines 3872-3874).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	OrgelArchivSchmidtCrawler.transform, line 3734
manual_count	manual_count	OrgelArchivSchmidtCrawler.transform, line 3733
stops_by_division	stops_by_division	OrgelArchivSchmidtCrawler.transform, line 3739
historical_specifications	disposition_versions[1:]	OrgelArchivSchmidtCrawler.transform, line 3741
raw_specification	current_spec	OrgelArchivSchmidtCrawler.transform, line 3737
country	country	OrgelArchivSchmidtCrawler.transform, line 3724

8. Research use, citation and rights

The website marks content as protected; no general open licence was established.[85,86,87]

S031. Orgeldatenbank Bayern

Catalogue key: orgeldb_bayern | Role: registered source | Documentary unit: Historical database record, including successive descriptions

1. Identity, scope and evidence

Source account. The portal serves an authorised copy of the 2009 database, described as exceeding 30,000 records with roughly 10,000 stoplists. It adds location matching using Wikipedia, Wikidata and OpenStreetMap and explicitly warns that some instruments labelled extant no longer survive in that form.[89,90]

Attribution recorded in the catalogue: Michael Bernhard (editor); Gesellschaft für Bayerische Musikgeschichte e.V. (original publisher). Scope: Bavaria; version 5, 2009.

2. Native record structure and contents

Selected document; 2026-09-11. CGI database record combining historical text fields with later geographic links[90]

- Postal place, district, administrative region, church and street address.
- Coordinates and Wikidata/OpenStreetMap references are displayed.
- Named builder associated with a work entry.
- New-build year and retained/extant status accompany the record.
- Stops are written inline by division.
- Manual divisions and a manual/register summary are present.
- Windchest and mechanical-action description.
- Case style and couplers/playing aids are stated.
- Historical retained/extant label is displayed.
- Bibliographic source with year and page plus source record number.
- Wikidata, Wikipedia and map links accompany the record.
- The source version is identified as 2009.

Repeated structures: Compact work statement, manual/register summary and inline division stoplists.

3. Discovery and acquisition

The adapter contains distinct paths for HTML detail pages and Access MDB input. The native record number, rather than a inferred physical-organ identity, supplies its primary documentary key.[5]

Configured source address: https://orgeldb.de. This is a registry setting, not a new availability test.

Configured primary identifier: orgeldb_bayern_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Deutschland" (aggregator/static_html_crawler.py, line 1417).
- default_region: "Bayern" (aggregator/static_html_crawler.py, line 1418).
- start_urls: ["https://orgeldb.de/cgi-bin/lkr"] (aggregator/static_html_crawler.py, line 1419).
- detail_url_patterns: ["dispo?satz="] (aggregator/static_html_crawler.py, line 1420).

4. Extraction and normalization

The HTML branch reads preformatted lines, location text, labelled builder and event information, technical counts and stop groups. It also extracts coordinates and a Wikidata identifier when present. The MDB branch is selected by a source-format marker and must be evaluated against its own fields. The HTML branch preserves status, source reference and record number among unstructured technical rows.[5]

5. Dates, relationships and historical states

Version 5 (2009) is the underlying documentary state; geographic enrichment is a separate layer.[90]

Database date and later geolocation are separate provenance layers. The record total cannot be read as a count of present organs. Treat apparent current-state statements as historical until corroborated.[89,90]

6. Preservation, loss and uncertainty

The two input formats should not be assumed to expose identical fields or versions. In the HTML branch, the derived description is truncated to 6,000 characters, while the original HTML is passed to the snapshot normalizer. A source number, Wikidata link and cross-source identity assertion have different evidential roles.[5]

Limit of the selected-document observation: The example URL identifier and displayed source record number differ. Neither should silently replace the other.

7. Implementation evidence and rule register

Registered class: OrgeldbBayernCrawler in aggregator/static_html_crawler.py; inheritance: OrgeldbBayernCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrgeldbBayernCrawler.crawl (lines 1422-1485).
- OrgeldbBayernCrawler.transform (lines 1614-1700).
- OrgeldbBayernCrawler._event_from_lines (lines 1866-1876).
- OrgeldbBayernCrawler._extract_links (lines 1940-1959).
- OrgeldbBayernCrawler._location_line (lines 1844-1853).
- OrgeldbBayernCrawler._manuals_stops (lines 1886-1888).
- OrgeldbBayernCrawler._mdb_event (lines 1791-1803).
- OrgeldbBayernCrawler._mdb_literature (lines 1805-1820).
- OrgeldbBayernCrawler._mdb_location (lines 1770-1789).
- OrgeldbBayernCrawler._parse_index_page (lines 1582-1612).

The JSON companion lists the remaining 2 source-specific routines and the other eligible inherited methods.

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	self._first_int(stop_count or raw_data.get('stop_count'))	OrgeldbBayernCrawler.transform, line 1663
manual_count	self._first_int(manual_count or raw_data.get('manual_count'))	OrgeldbBayernCrawler.transform, line 1662
stops_by_division	stops_by_division	OrgeldbBayernCrawler.transform, line 1664
dates	[parsed_year] if parsed_year else []	OrgeldbBayernCrawler._mdb_event, line 1799
country	self.default_country	OrgeldbBayernCrawler.transform, line 1643
raw_str	location_line or title	OrgeldbBayernCrawler.transform, line 1650

8. Research use, citation and rights

Permission to host the copy is stated; this is not a blanket permission for third-party redistribution.[89,90]

S032. Orgeldokumentationszentrum HSLU

Catalogue key: orgeldoc | Role: registered source | Documentary unit: Organ record linked to institutional documentation

1. Identity, scope and evidence

Source account. HSLU dates the Orgeldokumentationszentrum's activity as a competence centre to 2007-2025. Following Marco Brandazza's retirement in September 2025, the university maintains access to the organ database and catalogued archives. Its statement says that funding currently prevents additional inventory and publication projects.[91,92,93]

Attribution recorded in the catalogue: Hochschule Luzern; Zentral- und Hochschulbibliothek Luzern. Scope: Swiss organ culture and associated archival collections.

2. Native record structure and contents

Selected document; 2026-09-11. JSON detail endpoint with nested location, builder, organ and specification objects[93]

- Building, building part, locality and canton are populated.
- Builder, measure and year are separate attributes.
- New-build activity and year are populated.
- Register objects are nested under division objects.
- Manual and pedal range strings are supplied.
- Separate key and stop actions.
- Pitch with temperature and tuning-system fields.
- Bellows arrangement and wind-pressure value.
- Playing aids and a detailed comment are populated.
- A last-change date is supplied.

Repeated structures: Specification array contains divisions; each division contains register entries and comments.

3. Discovery and acquisition

The adapter assembles details, history, bibliography and media responses, with cached location context. Its detail model is distinct from the ordinary static-page family despite the common module.[5]

Configured source address: https://orgeldokumentationszentrum.ch. This is a registry setting, not a new availability test.

Configured primary identifier: orgeldoc_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Switzerland" (aggregator/static_html_crawler.py, line 2421).

4. Extraction and normalization

Location, organ, builder and disposition objects are read separately. Technical mappings include pipe count, key and stop action, wind pressure, tuning, pitch, manual and pedal compass, accessories and notes. Stop count is calculated from the extracted division entries. Manual count prefers a stated field and falls back to a count derived from divisions. Events, bibliography and media use their own helpers.[5]

5. Dates, relationships and historical states

Work year and last-change date are distinct keys.[93]

Institutional continuity of access must be distinguished from continued expansion. The physical archives, library holdings and online organ database are related but different research objects. Cite the relevant catalogue or record for individual holdings.[91,92,93]

6. Preservation, loss and uncertainty

Calculated stop totals are not independent source declarations; a division-derived manual count can differ from the physical console. History and principal builder information are merged conditionally. A scientific account must preserve the distinction between stated technical values, fallback context and calculated quantities.[5]

Limit of the selected-document observation: Some keys have null values. Register comments preserve materials, construction and compass-dependent exceptions.

7. Implementation evidence and rule register

Registered class: OrgelDocCrawler in aggregator/static_html_crawler.py; inheritance: OrgelDocCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrgelDocCrawler.crawl (lines 2423-2448).
- OrgelDocCrawler.transform (lines 2608-2686).
- OrgelDocCrawler._events (lines 2720-2736).
- OrgelDocCrawler._literature (lines 2784-2794).
- OrgelDocCrawler._media (lines 2775-2782).
- OrgelDocCrawler._refs_from_geo_locations (lines 2455-2457).
- OrgelDocCrawler._stops_by_division (lines 2748-2770).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	sum((len(stops) for stops in stops_by_division.values()))	OrgelDocCrawler.transform, line 2648
manual_count	self._first_int(orgel.get('anzmanuale')) or self._manual_count(stops_by_division)	OrgelDocCrawler.transform, line 2647
pipe_count	self._first_int(orgel.get('anzpfeifen'))	OrgelDocCrawler.transform, line 2649
stops_by_division	stops_by_division	OrgelDocCrawler.transform, line 2662
raw_specification	raw_spec or self._clean_text(disposition.get('textdisposition'))	OrgelDocCrawler.transform, line 2660
manual_compass	self._clean_text(orgel.get('umfangmanuale'))	OrgelDocCrawler.transform, line 2655

8. Research use, citation and rights

Continued access is promised; a general redistribution licence for the database and archives was not established.[91,92,93]

S033. Orgeldatenbank Sachsen

Catalogue key: orgelsachsen | Role: registered source | Documentary unit: Regional organ inventory record

1. Identity, scope and evidence

Source account. The source describes a project founded in autumn 2015 with Protestant and Catholic church partners. It estimates approximately 3,500 known organs, including lost instruments, and about 2,500 surviving organs in Saxony.[94,95]

Attribution recorded in the catalogue: Hochschule für Musik Carl Maria von Weber Dresden. Scope: Present-day Saxony, including historical instruments no longer extant.

2. Native record structure and contents

Selected document; 2026-09-11. Regional article with labelled organ description and historical specifications[95]

- Street address and locality are given.
- A location-map link is supplied.
- Predecessor and successor builders are distinguished.
- Replacement and later redisposition are dated.
- Two historical configurations have separate stoplists.
- Manual/pedal divisions are explicit; exact key compasses are not given here.
- Couplers and ancillary controls are listed.
- The predecessor is explicitly described as no longer surviving.
- Bibliographic citation includes a page locator.

Repeated structures: Separate division tables for the later instrument and an earlier predecessor.

Not established in this document check: Windchest and action labels exist, but no values appear in the selected record.; Pitch label is present without a value.; Gallery and image-credit headings exist without a verified image item..

3. Discovery and acquisition

The source page is acquired as HTML. Headline, location section, labelled technical data and tables supply different parts of the transfer.[96]

Configured source address: https://www.orgelforum-sachsen.de. This is a registry setting, not a new availability test.

Configured primary identifier: orgelsachsen_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The adapter separates a source-qualified identifier from Orgbase numbering. The headline is split for town/building, and location paragraphs can refine the address. Technical fields supply counts; action, windchest, pitch and repair information can remain descriptive notes. Tables with manual or pedal headers supply division-based stop rows. History is also retained as descriptive text.[96]

5. Dates, relationships and historical states

The selected page contains an 1825 predecessor and a 1905 replacement, with a specification after a 1971 alteration.[95]

These estimates describe the regional population, not a verified number of records exposed by the portal. Its territorial scope also differs from the church-owned instruments addressed by ORKASA, despite overlap between the two resources.[94,95]

6. Preservation, loss and uncertainty

A page may present historical specifications in addition to its principal disposition. Grouping table columns by division name does not itself provide a chronological configuration model. Notes retained as text remain valuable evidence even when no canonical field is populated. Default region and inferred title segments should be distinguished from explicit source labels.[96]

Limit of the selected-document observation: Several technical labels are blank. A field heading is not a populated measurement.

7. Implementation evidence and rule register

Registered class: OrgSachsenCrawler in aggregator/orgelsachsen_crawler.py; inheritance: OrgSachsenCrawler > BaseCrawler.[96]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrgSachsenCrawler.crawl (lines 40-165).
- OrgSachsenCrawler.transform (lines 295-605).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	tech_data['register']	OrgSachsenCrawler.transform, line 450
manual_count	tech_data['manuale']	OrgSachsenCrawler.transform, line 448
dates	[year] if year else []	OrgSachsenCrawler.transform, line 428
country	'Germany'	OrgSachsenCrawler.transform, line 338
raw_str	''	OrgSachsenCrawler.transform, line 343

8. Research use, citation and rights

The university project displays copyright; no general open licence was established.[94,95]

S034. Die Orgelseite

Catalogue key: orgelseite | Role: registered source | Documentary unit: Organ profile with photographs and specification

1. Identity, scope and evidence

Source account. The home page reported 6,622 organs in 83 countries and an update date of 5 September 2026. The imprint distinguishes public access from subscription and music-download areas and acknowledges photographs supplied by organ builders and other contributors.[97,98,99]

Attribution recorded in the catalogue: Die Orgelseite. Scope: International.

2. Native record structure and contents

Selected document; 2026-09-11. Filterable multi-record HTML table with links to specifications, photographs and music[99]

- Country, region, city and venue are populated in the organ table.
- Rows link to organ locations on maps.
- Builder names and year ranges are table columns.
- Build-year ranges represent documentary chronology, without an explicit event model in this table.
- Manual counts are tabulated.
- Rows identify publications, direct contacts and other websites.
- Rows link to specification, gallery and external source pages.
- Information stand and last-update are offered as separate optional columns.

Repeated structures: One table row per listed organ, with optional columns for ranks, stops, pipes and documentary sources.

Not established in this document check: Specification links are present; linked stoplist files were not opened.; Image and gallery links are present; their payloads were not inspected.; CD and MP3 columns link to the separate music catalogue..

3. Discovery and acquisition

A catalogue row, its disposition page and any acquired disposition workbook are distinct inputs. The source number is qualified as an Orgelseite identifier. The transform retains the catalogue cells and records which detail representation supplies the specification.[100]

Configured source address: https://die-orgelseite.de. This is a registry setting, not a new availability test.

Configured primary identifier: orgelseite_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

Location and initial builder attribution come from named catalogue columns. Both HTML and spreadsheet specification parsers are called; the spreadsheet result is selected when its parser reports a complete stoplist, otherwise the HTML result is selected. Divisions, stops, counts and raw specification text follow that choice. Photograph, gallery, audio and map links remain distinct. Non-empty catalogue cells are retained in a source-specific extension.[100]

5. Dates, relationships and historical states

Build-year ranges, information stand and last-update fields are distinct. Historical stoplists are counted separately.[99]

These totals are publisher statements. They do not measure data completeness or unrestricted access. The operator's name was not present in the retrieved imprint text, so the citation uses the publication title.[97,98,99]

6. Preservation, loss and uncertainty

The completeStopList flag is the parser's criterion, not independent confirmation of exhaustive organ documentation. A map link does not become a measured coordinate. A listing builder/year pair is projected as construction, so later interventions or abbreviated listings need documentary review. Workbook and HTML states should not be presumed contemporaneous.[100]

Limit of the selected-document observation: Displayed row number is an ordering position, not an established persistent organ identifier. Extensions, transmissions and digital additions are separate columns.

7. Implementation evidence and rule register

Registered class: OrgelseiteCrawler in aggregator/orgelseite_crawler.py; inheritance: OrgelseiteCrawler > BaseCrawler.[100]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrgelseiteCrawler.crawl (lines 527-657).
- OrgelseiteCrawler.transform (lines 659-798).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	detail_specification['stop_count']	OrgelseiteCrawler.transform, line 750
manual_count	detail_specification['manual_count']	OrgelseiteCrawler.transform, line 749
stops_by_division	detail_specification['stops_by_division']	OrgelseiteCrawler.transform, line 753
dates	[raw_data.get('Baujahr', '')]	OrgelseiteCrawler.transform, line 732
country	self.country_mapping.get(raw_data.get('Land', ''), raw_data.get('Land', ''))	OrgelseiteCrawler.transform, line 720

8. Research use, citation and rights

Text and photographs are described as copyrighted; contributed photographs may retain third-party rights.[97,98,99]

S035. Orgelsite NL

Catalogue key: orgelsite | Role: registered source | Documentary unit: Organ webpage with photographs and search tags

1. Identity, scope and evidence

Source account. The site reports approximately 35,000 organs and offers searches by place, builder, year and manual count. It accepts contributed photographs and explicitly reserves copyright in photographs, with exceptions identified for Wikimedia and YouTube material.[101,102]

Attribution recorded in the catalogue: de Orgelsite / orgelsite.nl. Scope: International; Dutch, German and English introductory text.

2. Native record structure and contents

Selected document; 2026-09-11. Tagged organ article with contributed photography, historical prose and specification[102]

- Place, church and region are identified.
- Successive builders and restorers are named.
- Construction, replacement and restoration chronology.
- Stops are grouped in division paragraphs.
- Division ranges and note counts are supplied.
- Mechanical action and slider chests are described.
- A dated re-leathering of reservoir bellows is recorded.
- Case alteration and added performance controls are described.
- Surviving older case is distinguished from the replacement instrument.
- Multiple photographs have named authors and dates.
- The specification is attributed to a named contributor and website.
- Contributor website and taxonomic tags are linked.
- Page author, publication/update dates and photo credits are separate.

Repeated structures: Narrative chronology followed by division-based stop paragraphs and coupling list.

3. Discovery and acquisition

The subclass accepts single-segment detail paths on its configured host and excludes category, tag, author, feed and media paths. This discovery rule is narrower than all pages on the website.[5]

Configured source address: https://www.orgelsite.nl. This is a registry setting, not a new availability test.

Configured primary identifier: orgelsite_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- start_urls: ["https://www.orgelsite.nl/"] (aggregator/static_html_crawler.py, line 1003).
- exclude_url_patterns: ["/categorie/", "/tag/", "/author/", "/feed/", "/wp-content/", "mailto:", "#"] (aggregator/static_html_crawler.py, line 1004).

4. Extraction and normalization

The transform uses taxonomy tags to identify a year, builder label and certain country labels. It suppresses implausible cached stop counts, retains the tags and records the parsed stop count. It otherwise uses the static HTML metadata, text and stoplist routines.[5]

5. Dates, relationships and historical states

Historical instrument years, photo years and publication/update dates are different attributes.[102]

The approximate total is not an independently enumerated corpus. Tags aid discovery but should not be mistaken for controlled authority records. Photograph authorship needs item-level treatment.[101,102]

6. Preservation, loss and uncertainty

Tags may describe classification rather than a complete chronology. A calculated number of parsed rows must not be cited as a source-declared total. Detail discovery and tag vocabulary are code assumptions that may miss valid pages or differently labelled information.[5]

Limit of the selected-document observation: The named author of the specification can differ from the page author and photographers.

7. Implementation evidence and rule register

Registered class: OrgelsiteNlCrawler in aggregator/static_html_crawler.py; inheritance: OrgelsiteNlCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- OrgelsiteNlCrawler.transform (lines 1023-1049).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	''	OrgelsiteNlCrawler.transform, line 1035
country	country	OrgelsiteNlCrawler.transform, line 1048
specifications (read)	transformed.get('specifications', {})	OrgelsiteNlCrawler.transform, line 1038
stops (read)	specs.get('stops', [])	OrgelsiteNlCrawler.transform, line 1039
html (read)	raw_data.get('html')	OrgelsiteNlCrawler.transform, line 1024
locations (read)	transformed.get('locations')	OrgelsiteNlCrawler.transform, line 1044

8. Research use, citation and rights

The page requires prior written permission from the webmaster and photographer for republication of copyrighted photographs.[101,102]

S036. Orgelverzeichnis Schweiz / Liechtenstein

Catalogue key: orgelverzeichnis_ch | Role: registered source | Documentary unit: Organ profile with dated specification

1. Identity, scope and evidence

Source account. The rights statement credits information to congregations, builders, archives and contributors, with much fieldwork undertaken by Peter Fasler and François Comment. Individual profiles identify their information or image suppliers.[103,104,105]

Attribution recorded in the catalogue: Peter Fasler; François Comment / Orgelverzeichnis Schweiz. Scope: Switzerland and Liechtenstein.

2. Native record structure and contents

Selected document; 2026-09-11. Compact static profile with technical summary, specification and dated narrative[105]

- Institution, locality and canton are given.
- Builder and build year are populated.
- The construction is described in a dated note.
- Three named stops with pitches are tabulated.
- Manual compass is given.
- Separate action and windchest attributes.
- A transposition mechanism is described with two reference pitches.
- Retractable keyboard and transposition device are listed.
- Photograph is credited to an organ builder.
- Information supplier is named.
- Links connect regional lists, builders and HSLU.
- Information reference date and separate information/image credits.

Repeated structures: Label-value technical fields followed by stop rows and playing aids.

3. Discovery and acquisition

The index parser reads links and layout-dependent row context, including positioning information. The transform can override town, building and region with the acquired index values.[5]

Configured source address: https://orgelverzeichnis.ch. This is a registry setting, not a new availability test.

Configured primary identifier: orgelverzeichnis_ch_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Switzerland" (aggregator/static_html_crawler.py, line 3428).
- start_urls: ["https://orgelverzeichnis.ch/Orgelverzeichnis/auswahl_kt.htm"] (aggregator/static_html_crawler.py, line 3429).
- detail_url_patterns: ["profile/"] (aggregator/static_html_crawler.py, line 3430).
- exclude_url_patterns: ["index.htm", "auswahl", "instr", "urheber", "quellen", "dank", ".pdf"] (aggregator/static_html_crawler.py, line 3431).

4. Extraction and normalization

Detail extraction otherwise inherits the common static-page implementation. The index and page are therefore two evidential layers. Later source-specific Swiss date and specification-account helpers are separate from this inherited transform and must be cited separately when their results are used.[5]

5. Dates, relationships and historical states

Construction year and documentary stand date are separate.[105]

Preserve the profile's state date and the contributor attached to a specification. The legacy MAGIX address in Aggregator's investigation audit is an earlier location, not a second independent source.[103,104,105]

6. Preservation, loss and uncertainty

An information-reference date belongs to the specification account it labels, not automatically the instrument's construction. Layout-derived rows may fail under a changed presentation. The initial transfer should not be credited with all later offline evidence recovery.[5]

Limit of the selected-document observation: Supplier credits for information and images are distinct; the stand date limits claims of currency.

7. Implementation evidence and rule register

Registered class: OrgelverzeichnisChCrawler in aggregator/static_html_crawler.py; inheritance: OrgelverzeichnisChCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrgelverzeichnisChCrawler.crawl (lines 3439-3497).
- OrgelverzeichnisChCrawler.transform (lines 3539-3548).
- OrgelverzeichnisChCrawler._location_parts (lines 3593-3600).
- OrgelverzeichnisChCrawler._parse_index_page (lines 3508-3537).
- OrgelverzeichnisChCrawler._stops_from_row (lines 3587-3591).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
town (read)	raw_data.get('town')	OrgelverzeichnisChCrawler.transform, line 3542
building (read)	raw_data.get('building')	OrgelverzeichnisChCrawler.transform, line 3544
region (read)	raw_data.get('region')	OrgelverzeichnisChCrawler.transform, line 3546
locations (read)	transformed.get('locations')	OrgelverzeichnisChCrawler.transform, line 3541
timeout (read)	self.config.get('timeout', 30)	OrgelverzeichnisChCrawler._fetch_html, line 3434
task_id (read)	kwargs.get('task_id')	OrgelverzeichnisChCrawler.crawl, line 3440

8. Research use, citation and rights

The policy permits specified non-commercial uses, requires visible source and item credits, and requires permission for commercial use. It also preserves rights in contributed images.[103,104,105]

S037. OrguesCH

Catalogue key: orguesch | Role: registered source | Documentary unit: Project or instrument presentation

1. Identity, scope and evidence

Source account. The current home page describes organ consultancy and coordination and presents nearly 60 instruments built, restored or revised with partner builders. It also advertises other activities, including Allen Organs representation.[106,107]

Attribution recorded in the catalogue: Orgues.ch. Scope: Selected Swiss construction, restoration and revision projects.

2. Native record structure and contents

Selected document; 2026-09-11. [107]

Access finding. The inherited detail route returned an organ-not-found message despite HTTP 200. This does not establish that the portfolio or instrument has disappeared. No native detail structure is inferred from this failed check.

3. Discovery and acquisition

Index cards are selected through organ-detail URL patterns. Titles are sought in headings, image alternative text, styled paragraphs and finally link text; the last paragraph can supply town context.[5]

Configured source address: https://www.orgues.ch. This is a registry setting, not a new availability test.

Configured primary identifier: orguesch_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Switzerland" (aggregator/static_html_crawler.py, line 1057).
- start_urls: ["https://www.orgues.ch/orgues"] (aggregator/static_html_crawler.py, line 1058).
- detail_url_patterns: ["/orgues/orgue-"] (aggregator/static_html_crawler.py, line 1059).

4. Extraction and normalization

The subclass supplies card discovery and a Switzerland default. Detailed extraction otherwise inherits the static HTML transform. That shared route preserves main text, metadata, table rows, images and links while attempting a labelled specification and a builder/year summary.[5]

5. Dates, relationships and historical states

This is a selected project portfolio rather than a Swiss national inventory. Its commercial and collaborative context belongs in the source description. The current presentation should not inherit the broader implications of an older generic database label.[106,107]

6. Preservation, loss and uncertainty

The presence of a structured card does not imply that the detail page has the same metadata model. Country defaults and inherited first-year/construction logic are interpretive operations. Source text retained outside structured fields should be examined before declaring information absent.[5]

7. Implementation evidence and rule register

Registered class: OrguesChCrawler in aggregator/static_html_crawler.py; inheritance: OrguesChCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].

- StaticHtmlOrganCrawler.transform (lines 434-516) [inherited].
- OrguesChCrawler._parse_index_page (lines 1061-1076).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	StaticHtmlOrganCrawler.transform, line 464
manual_count	manual_count	StaticHtmlOrganCrawler.transform, line 463
rank_count	rank_count	StaticHtmlOrganCrawler.transform, line 465
stops_by_division	stops_by_division	StaticHtmlOrganCrawler.transform, line 466
country	raw_data.get('country') or self.default_country	StaticHtmlOrganCrawler.transform, line 453
raw_str	title	StaticHtmlOrganCrawler.transform, line 457

8. Research use, citation and rights

The current page states that all rights are reserved.[106,107]

S038. Orgues France

Catalogue key: orguesfrance | Role: registered source | Documentary unit: Compiled organ profile

1. Identity, scope and evidence

Source account. The personal site's introduction says that it has collected information from the web since 2008. On 10 September 2026 it reported more than 3,700 pipe organs in 38 countries. It also offers multilingual stop-name and organ-term aids.[108,109]

Attribution recorded in the catalogue: Orgues en France et dans le monde. Scope: France and other countries.

2. Native record structure and contents

Selected document; 2026-09-11. Static compiled organ profile with metadata block, narrative and linked specification/media[109]

- Country, region, department, city and venue are labelled.
- Builder and construction year are labelled.
- A dated account covers predecessors and the new concert organ.
- Different actions of the two consoles are described.
- Variable wind pressure and high-pressure stops are discussed.
- Two-console arrangement and performance controls are described.
- Map, photograph, specification and video links are supplied.
- Page creation and update dates are explicit.

Repeated structures: Geographic/building/builder fields precede history and technical material.

Not established in this document check: External video presentations are linked; their contents were not inspected.; Specification is offered as linked material; a text stoplist was not observed in this response.; Photograph links are present; their contents were not separately inspected..

3. Discovery and acquisition

The subclass implements source-specific crawling and index parsing before passing detail pages to the inherited transform.[5]

Configured source address: http://orguesfrance.com. This is a registry setting, not a new availability test.

Configured primary identifier: orguesfrance_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "France" (aggregator/static_html_crawler.py, line 1975).
- start_urls: ["http://orguesfrance.com/Select.php"] (aggregator/static_html_crawler.py, line 1976).
- detail_url_patterns: ["orguesfrance.com/"] (aggregator/static_html_crawler.py, line 1977).
- exclude_url_patterns: ["select", "index.html", ".jpg", ".png", ".gif"] (aggregator/static_html_crawler.py, line 1978).

4. Extraction and normalization

The inherited implementation collects main text, label/value metadata, selected stoplist text, table rows, images and links. It projects a builder/year summary and source-reference context. These are implemented operations; they do not establish a formal native database schema behind the website.[5]

5. Dates, relationships and historical states

Page creation/update dates are separate from instrument construction and relocation events.[109]

Its explicit compilation from other websites makes source dependence a central issue. Agreement between compiled websites does not, by itself, establish independent corroboration; the cited antecedents must be examined. The two Orguesfrance entries in the investigation audit are one registered source.[108,109]

6. Preservation, loss and uncertainty

The configured crawl route bounds what is discovered. A narrative can contain several configurations even when the output has one principal specification. Generic label and date heuristics require a record-level check before their results support historical comparisons.[5]

Limit of the selected-document observation: The profile discusses predecessor instruments in the same institution. Media and specification may be separate linked assets.

7. Implementation evidence and rule register

Registered class: OrguesFranceCrawler in aggregator/static_html_crawler.py; inheritance: OrguesFranceCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- OrguesFranceCrawler.crawl (lines 1980-2014).
- StaticHtmlOrganCrawler.transform (lines 434-516) [inherited].
- OrguesFranceCrawler._parse_index_page (lines 2034-2066).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	StaticHtmlOrganCrawler.transform, line 464
manual_count	manual_count	StaticHtmlOrganCrawler.transform, line 463
rank_count	rank_count	StaticHtmlOrganCrawler.transform, line 465
stops_by_division	stops_by_division	StaticHtmlOrganCrawler.transform, line 466
country	raw_data.get('country') or self.default_country	StaticHtmlOrganCrawler.transform, line 453
raw_str	title	StaticHtmlOrganCrawler.transform, line 457

8. Research use, citation and rights

No general open licence was established from the consulted home page.[108,109]

S039. OSWA Pipe Organs in Western Australia

Catalogue key: oswa | Role: registered source | Documentary unit: Regional organ profile

1. Identity, scope and evidence

Source account. The society describes preservation and advisory work on Western Australia's organ heritage and links to its pipe-organ directory. Its wider remit also includes electronic and hybrid instruments.[110,111]

Attribution recorded in the catalogue: Organ Society of Western Australia. Scope: Western Australia.

2. Native record structure and contents

Selected document; 2026-09-11. Gazetteer form with property, previous-organ, present-organ and document-control blocks[111]

- Institution, street, city, state, country and room fields.
- Named maker and completion/installation year.
- Original installation, relocation and later work are separately described.
- Division stop rows with pitch and total stops/ranks/pipes.
- Divisions and manual count are populated; key-compass fields are blank.
- Key and stop actions are separately populated.
- Placement, open casework and console configuration are supplied.
- The organ is described as in use and in good condition.
- Photographs have named contributors.
- Gazetteer editions are cited in document control.
- Comparable instrument and builder links.
- Document control identifies original entries, contributor and revision dates.

Repeated structures: Label-value fields, division stop rows and repeated dated work narratives.

Not established in this document check: A wind-pressure label is present without a value..

3. Discovery and acquisition

The town index is parsed as a table: town, building, builder, date, manual, rank and stop counts occupy positional cells; an optional additional cell retains pipes. Detail pages are fetched separately.[5]

Configured source address: https://www.oswa.org.au. This is a registry setting, not a new availability test.

Configured primary identifier: oswa_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Australia" (aggregator/static_html_crawler.py, line 917).
- default_region: "Western Australia" (aggregator/static_html_crawler.py, line 918).
- start_urls: ["https://www.oswa.org.au/sup/WAOrgans/Indexes/WAOrganTownIndex.html"] (aggregator/static_html_crawler.py, line 919).
- detail_url_patterns: ["/sup/waorgans/"] (aggregator/static_html_crawler.py, line 920).
- exclude_url_patterns: ["/indexes/"] (aggregator/static_html_crawler.py, line 921).

4. Extraction and normalization

The detail override prefers explicitly labelled builder, completion/installation date and total-stop fields. Specification extraction chooses the last leaf-table candidate containing a recognised division line. Declared, parsed and reference stop counts are kept separately, with declared count preferred for the main value.[5]

5. Dates, relationships and historical states

Previous organs, current organ, relocation dates and document-control dates are separate.[111]

Aggregator's pipe-organ selection should be distinguished from the society's complete activities. State-level profiles can complement OHTA, but overlapping coverage must be checked at the instrument and documentary-source level.[110,111]

6. Preservation, loss and uncertainty

Choosing the last candidate is an extraction rule, not independent evidence that it describes the latest organ. Parsed-row counts and declared stops are different quantities. Positional index columns must be checked when the source layout changes.[5]

Limit of the selected-document observation: Many form labels are unfilled. The existence of a wind-pressure or heritage-number field does not establish a value.

7. Implementation evidence and rule register

Registered class: OswaCrawler in aggregator/static_html_crawler.py; inheritance: OswaCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- OswaCrawler.transform (lines 936-959).
- OswaCrawler._extract_raw_spec (lines 923-934).
- OswaCrawler._parse_index_page (lines 961-995).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	exact_fields.get('total number of stops') or prepared.get('stop_count', '')	OswaCrawler.transform, line 949
specifications (read)	transformed.get('specifications', {})	OswaCrawler.transform, line 951
metadata (read)	transformed.get('extensions', {}).get(self.source_key, {}).get('metadata', {})	OswaCrawler.transform, line 952
builder's name (read)	exact_fields.get("builder's name")	OswaCrawler.transform, line 947
builder (read)	prepared.get('builder', '')	OswaCrawler.transform, line 947
date of completion/installation (read)	exact_fields.get('date of completion/installation')	OswaCrawler.transform, line 948

8. Research use, citation and rights

The society displays an all-rights-reserved notice; no general reuse licence was established.[110,111]

S040. PCorgan Hauptwerk Sample Set Registry

Catalogue key: pcorgan_hauptwerk_samplesets | Role: registered source | Documentary unit: Comparative sample-set table row

1. Identity, scope and evidence

Source account. The comparison table records product, producer, style, size, indicative price and technical characteristics. It defines its memory column as a particular Hauptwerk-plus-set loading scenario. The author explains the site's church-organ focus and discloses receiving sets for review or beta testing.[112,113]

Attribution recorded in the catalogue: Gert van Ginkel / PCorgan. Scope: Mainly church-organ sample sets for Hauptwerk.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects catalogue list links and links matching the SampleSets pattern. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://www.pcorgan.com/SampleSetsEN.html. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "registry" (aggregator/grandorgue_library_crawler.py, line 1657).

4. Extraction and normalization

The implemented unit is third-party sample-set registry items. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

The visible prices are explicitly indicative and may be outdated. Memory values are not interchangeable with download sizes or modern maximum-memory configurations. This source is a comparative index, not the producer's release record.[112,113]

6. Preservation, loss and uncertainty

A registry may summarise producer material; current technical details and rights require the cited producer edition. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: PcorganHauptwerkSamplesetsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: PcorganHauptwerkSamplesetsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- PcorganHauptwerkSamplesetsCrawler.parse_html (lines 1659-1701).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'registry_entry'	PcorganHauptwerkSamplesetsCrawler.parse_html, line 1686
requires_sample_set	False	PcorganHauptwerkSamplesetsCrawler.parse_html, line 1688
is_odf_only	False	PcorganHauptwerkSamplesetsCrawler.parse_html, line 1687

8. Research use, citation and rights

Copyright is asserted; permissions for linked audio and products remain with their respective sources.[112,113]

S041. Piotr Grabowski GrandOrgue Versions

Catalogue key: piotr_grabowski_go_versions | Role: registered source | Documentary unit: GrandOrgue conversion or compatibility record

1. Identity, scope and evidence

Source account. The page documents installation, patching and feature differences between GrandOrgue and Hauptwerk versions. It states that some features are absent and that presentation videos were recorded using Hauptwerk.[114]

Attribution recorded in the catalogue: Piotr Grabowski. Scope: GrandOrgue versions of the producer's sample sets.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects article, list, paragraph and table nodes, with download-link fallback. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://piotrgrabowski.pl/grandorgue-versions/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_downloads" (aggregator/grandorgue_library_crawler.py, line 884).

4. Extraction and normalization

The implemented unit is GrandOrgue versions and their linked assets. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Shared recordings do not imply identical software behaviour. Link a GrandOrgue version to the corresponding product while preserving format, release and known limitations. This key and the instruments catalogue share a producer and much underlying material.[114]

6. Preservation, loss and uncertainty

A port or format version should remain linked to its underlying product without collapsing edition and format into one identity. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: PiotrGrabowskiGoVersionsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: PiotrGrabowskiGoVersionsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- PiotrGrabowskiGoVersionsCrawler.parse_html (lines 886-959).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'odf'	PiotrGrabowskiGoVersionsCrawler.parse_html, line 932
requires_sample_set	True	PiotrGrabowskiGoVersionsCrawler.parse_html, line 936
is_odf_only	True	PiotrGrabowskiGoVersionsCrawler.parse_html, line 938
producer_label	'Piotr Grabowski'	PiotrGrabowskiGoVersionsCrawler.parse_html, line 925

8. Research use, citation and rights

Software conversion information does not establish permission to redistribute the underlying recordings.[114]

S042. Piotr Grabowski Hauptwerk Instruments

Catalogue key: piotr_grabowski_instruments | Role: registered source | Documentary unit: Instrument-based sample-set product

1. Identity, scope and evidence

Source account. Catalogue cards expose builder, date, stops and keyboard counts. Free products and unreleased entries are marked separately. The list also contains a harmonium, so the producer's instrument catalogue is broader than pipe organs alone.[115]

Attribution recorded in the catalogue: Piotr Grabowski. Scope: International product catalogue with paid, free and unreleased entries.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects links to instrument pages on the producer host. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://piotrgrabowski.pl/instruments/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_catalog" (aggregator/grandorgue_library_crawler.py, line 1269).

4. Extraction and normalization

The implemented unit is Hauptwerk instrument products. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Preserve release status and instrument type. The card's construction date belongs to the represented instrument, not automatically to the sample set. Duplicate discovery through other registries should resolve to a product record rather than another physical organ.[115]

6. Preservation, loss and uncertainty

The same represented organ can have several products or format versions; catalogue identity must retain that distinction. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: PiotrGrabowskiInstrumentsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: PiotrGrabowskiInstrumentsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- PiotrGrabowskiInstrumentsCrawler.parse_html (lines 1271-1329).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stops	PiotrGrabowskiInstrumentsCrawler.parse_html, line 1316
manual_count	_first_int(keyboards)	PiotrGrabowskiInstrumentsCrawler.parse_html, line 1314
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'sample_set'	PiotrGrabowskiInstrumentsCrawler.parse_html, line 1304
requires_sample_set	False	PiotrGrabowskiInstrumentsCrawler.parse_html, line 1306
is_odf_only	False	PiotrGrabowskiInstrumentsCrawler.parse_html, line 1305

8. Research use, citation and rights

The producer asserts copyright; free products still require their own licence check.[115]

S043. Pipedreams Organ Gallery

Catalogue key: pipedreams | Role: registered source | Documentary unit: Organ-gallery profile linked to programme and performance material

1. Identity, scope and evidence

Source account. The live site is organised around programmes and performances, with dated episodes and related material. Aggregator selects organ-gallery information from that broader broadcasting resource.[116,1,117]

Attribution recorded in the catalogue: Pipedreams. Scope: Instruments documented in a broadcasting context.

2. Native record structure and contents

Selected document; 2026-09-11. Organ-gallery hub connecting an instrument to broadcast programmes and external documents[117]

- Concert venue, city and country are in the title.
- Manufacturer and construction year are in the title.
- A manufacturer PDF and an external organ account are linked.
- Venue, manufacturer and organ-description links are present.

Repeated structures: Related-document links followed by many programme cards.

Not established in this document check: Specification is linked as an external manufacturer PDF, not transcribed in the inspected page.; Programme cards establish a route to broadcasts; audio files were not inspected..

3. Discovery and acquisition

The crawler interprets embedded application data and profile structures, with separate discovery logic for gallery pages.[5]

Configured source address: https://www.pipedreams.org. This is a registry setting, not a new availability test.

Configured primary identifier: pipedreams_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- start_urls: ["https://www.pipedreams.org/page/organ-gallery"] (aggregator/static_html_crawler.py, line 2966).
- detail_url_patterns: ["/profile/"] (aggregator/static_html_crawler.py, line 2967).

4. Extraction and normalization

Profile title and rich body supply identification and description. Image resources, related links and audio contribution results have specialised handlers. Publication date, profile ID, resource type and canonical slug are retained. The transform intentionally leaves the raw specification empty and adds a source-scope note.[5]

5. Dates, relationships and historical states

The instrument construction year in the title differs from broadcast episode dates.[117]

The recording, programme, performer and instrument are different entities. A programme's publication date cannot be assigned as a date for an organ's construction or documented specification. The inspected homepage does not establish an exhaustive inventory of organs.[116,1,117]

6. Preservation, loss and uncertainty

That empty specification is an output decision, not an audit of every gallery page. Publication date concerns the profile, not organ construction. A title-derived builder/year is less explicit than a labelled source field; media and broadcast associations should not be treated as organ measurements.[5]

Limit of the selected-document observation: A gallery page primarily aggregates performance context. Linked specifications were not downloaded.

7. Implementation evidence and rule register

Registered class: PipedreamsCrawler in aggregator/static_html_crawler.py; inheritance: PipedreamsCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- PipedreamsCrawler.crawl (lines 2998-3056).
- PipedreamsCrawler.transform (lines 3096-3162).
- PipedreamsCrawler._location_from_profile_url (lines 3276-3283).
- PipedreamsCrawler._parse_index_page (lines 3068-3094).
- PipedreamsCrawler._parse_profile_title (lines 3245-3259).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
raw_specification	''	PipedreamsCrawler.transform, line 3130
country	country	PipedreamsCrawler.transform, line 3121
raw_str	title	PipedreamsCrawler.transform, line 3125
html (read)	raw_data.get('html')	PipedreamsCrawler.transform, line 3097
raw_html (read)	raw_data.get('raw_html')	PipedreamsCrawler.transform, line 3097
town (read)	raw_data.get('town')	PipedreamsCrawler.transform, line 3104

8. Research use, citation and rights

Streaming or linking to a programme does not establish permission to redistribute its audio.[116,1,117]

S044. Pipe Organ Map

Catalogue key: pipeorganmap | Role: registered source | Documentary unit: Organ record linked to builders and sample-set creators

1. Identity, scope and evidence

Source account. The site combines map-based discovery with organ, builder and sample-set-creator records and links to API documentation. Its footer identifies Leif-Erik Hallmann and labels content CC BY-SA 4.0.[118,119]

Attribution recorded in the catalogue: Leif-Erik Hallmann / Pipe Organ Map. Scope: International map and database.

2. Native record structure and contents

Selected document; 2026-09-11. Linked organ profile with metadata, division stop table, map and source links[119]

- Street address, city and country are supplied.
- Map and distance-based neighbouring instruments are displayed.
- Builder is a linked entity.
- Dated narrative distinguishes installation and later relocations.
- Named divisions contain stop names and pitches.
- Division count and manual/pedal grouping are supplied.
- Key action is a labelled attribute; prose also describes stop action.
- Tuning is a labelled attribute.
- Case and console arrangement are described in prose.
- The record cites Musique et musiciens and INdO.
- Related instrument, builder and upstream record links are explicit.
- A version-history route is exposed.

Repeated structures: Stop divisions, nearby instruments, same-builder instruments and similarity suggestions.

Not established in this document check: The selected record explicitly states that no video/audio samples are available..

3. Discovery and acquisition

The subclass extracts organ links from the map/index representation and provides an additional specification-selection routine.[5]

Configured source address: https://pipeorganmap.com. This is a registry setting, not a new availability test.

Configured primary identifier: pipeorganmap_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- start_urls: ["https://pipeorganmap.com/organs"] (aggregator/static_html_crawler.py, line 1279).
- detail_url_patterns: ["/organ/"] (aggregator/static_html_crawler.py, line 1280).
- exclude_url_patterns: ["/preview", ".pdf", "/liked", "/nearby"] (aggregator/static_html_crawler.py, line 1281).

4. Extraction and normalization

It uses the shared HTML transform after source-specific discovery and stoplist handling. Map/list context, textual description, division strings and images are different forms of evidence. Source identifiers derived from URLs identify documentary entries rather than proving unique physical instruments.[5]

5. Dates, relationships and historical states

Build year and narrative interventions coexist with a version-history link.[119]

The explicit separation of builders and sample-set creators supports cross-domain links. Map coordinates and popular-organ selections should not be treated as measures of historical coverage. Any claim imported from another source still needs its attribution.[118,119]

6. Preservation, loss and uncertainty

A map entry and a coordinate measurement are not interchangeable. The generic transform can retain text without resolving all fields, and its first-year fallback can be too broad for a narrative containing several dates. Geographic reach must be established from a dated collection inventory.[5]

Limit of the selected-document observation: Sample-set suggestions are recommendation links, not proof of sampled-instrument identity. Only the pipe-organ record is included here.

7. Implementation evidence and rule register

Registered class: PipeOrganMapCrawler in aggregator/static_html_crawler.py; inheritance: PipeOrganMapCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- StaticHtmlOrganCrawler.crawl (lines 78-129) [inherited].
- StaticHtmlOrganCrawler.transform (lines 434-516) [inherited].
- PipeOrganMapCrawler._extract_raw_spec (lines 1320-1336).
- PipeOrganMapCrawler._parse_index_page (lines 1283-1300).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	stop_count	StaticHtmlOrganCrawler.transform, line 464
manual_count	manual_count	StaticHtmlOrganCrawler.transform, line 463
rank_count	rank_count	StaticHtmlOrganCrawler.transform, line 465
stops_by_division	stops_by_division	StaticHtmlOrganCrawler.transform, line 466
country	raw_data.get('country') or self.default_country	StaticHtmlOrganCrawler.transform, line 453
raw_str	title	StaticHtmlOrganCrawler.transform, line 457

8. Research use, citation and rights

The footer states CC BY-SA 4.0 for content; third-party material requires its own credit and scope check.[118,119]

S045. RIdIM Music Iconography

Catalogue key: ridim_iconography | Role: registered source | Documentary unit: Music-iconography catalogue record

1. Identity, scope and evidence

Source account. The configured endpoint is the German RIdIM catalogue on musiconn, with basic and advanced searching. Its interface credits the Bayerische Staatsbibliothek and attaches specific licences to displayed collection images.[120,121]

Attribution recorded in the catalogue: RIdIM Germany / Bayerische Staatsbibliothek. Scope: Visual representations of music and instruments.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

Result and detail material is merged into an iconographic record; query and result context remain part of its acquisition history.[122,123]

Configured source address: https://ridim.musiconn.de/ridim/. This is a registry setting, not a new availability test.

Configured primary identifier: ridim_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The normalizer separates artwork, holding institution, iconographic subject matter, media and links. Native record identifiers, local object numbers and sigla have distinct roles. Fallback identity can be derived from identifying text. Original detail-field labels remain relevant to interpreting the constructed artwork and iconography fields.[122,123]

5. Dates, relationships and historical states

A depicted organ is not a surviving instrument record. Object identifier, holding institution, creator, date and depicted subject must remain distinct. The endpoint should not be cited as though it covered every international RIdIM database.[120,121]

6. Preservation, loss and uncertainty

An image depicting an organ is evidence about a depiction, not automatically about a surviving instrument. Production date, depicted event date and catalogue modification date must be distinguished. An image reference does not establish redistribution rights or physical access to the object.[122,123]

7. Implementation evidence and rule register

Registered class: RidimIconographyCrawler in aggregator/ridim_iconography_crawler.py; inheritance: RidimIconographyCrawler > BaseCrawler.[122,123]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- RidimIconographyCrawler.crawl (lines 94-156).
- RidimIconographyCrawler.transform (lines 284-300).
- RidimIconographyCrawler._respectful_pause (lines 544-555).
- RidimIconographyCrawler.parse_detail_html (lines 219-282).
- RidimIconographyCrawler.parse_rss (lines 165-209).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
raw_data_json (read)	payload.get('raw_data_json')	RidimIconographyCrawler.transform, line 286
crawl_run_id (read)	payload.get('crawl_run_id')	RidimIconographyCrawler.transform, line 298
transform_run_id (read)	payload.get('transform_run_id')	RidimIconographyCrawler.transform, line 299
raw_html (read)	payload.get('raw_html')	RidimIconographyCrawler.transform, line 297
task_id (read)	kwargs.get('task_id')	RidimIconographyCrawler.crawl, line 96
timeout (read)	self.config.get('timeout', 30)	RidimIconographyCrawler.fetch_detail, line 214

8. Research use, citation and rights

Displayed images carry different licences. A licence beside one illustration does not cover the entire catalogue.[120,121]

S046. Sakralorgelforum

Catalogue key: sakralorgelforum | Role: registered source | Documentary unit: Discussion thread and dated post

1. Identity, scope and evidence

Source account. The public index includes instrument and musical discussions alongside liturgy, community and marketplace sections. It exposes recent activity and topic counts, but these interface statistics do not establish a complete exportable corpus.[124]

Attribution recorded in the catalogue: Sakralorgelforum. Scope: Sacred music, organs, playing, liturgy and community discussion.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The subclass uses the forum-thread family with its source-specific configuration. The effective inherited methods are listed explicitly in this atlas.[44,45]

Configured source address: https://www.sakralorgelforum.net. This is a registry setting, not a new availability test.

Configured primary identifier: thread_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

The common transform combines acquired thread pages, preserves listing context and creates a thread record with nested posts. Its inheritance should be described directly rather than pretending that a separate class supplies an independent schema. Source-qualified thread/post identities remain distinct from people or instruments mentioned.[44,45]

5. Dates, relationships and historical states

Discussion is useful for discovery and reception research. Assertions about instruments need attribution to the particular post and corroboration appropriate to the question. Participant prose and personal labels are excluded from this package.[124]

6. Preservation, loss and uncertainty

The family's pagination and quote-attribution limitations also apply here. A successful common parser does not prove that every source-specific theme or historical layout is supported. Author profile and reaction metadata are documentary context, not an organological reliability score.[44,45]

7. Implementation evidence and rule register

Registered class: SakralorgelForumCrawler in aggregator/kirchenorgelforum_crawler.py; inheritance: SakralorgelForumCrawler > KirchenorgelForumCrawler > BaseCrawler.[44,45]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- KirchenorgelForumCrawler.crawl (lines 156-265) [inherited].
- KirchenorgelForumCrawler.transform (lines 340-394) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
html (read)	raw_data.get('html')	KirchenorgelForumCrawler.transform, line 341
raw_html (read)	raw_data.get('raw_html')	KirchenorgelForumCrawler.transform, line 341
url (read)	raw_data.get('url')	KirchenorgelForumCrawler.transform, line 345
forum_id (read)	forum.get('forum_id')	KirchenorgelForumCrawler.transform, line 358
forum_title (read)	forum.get('forum_title')	KirchenorgelForumCrawler.transform, line 359
raw_index_html (read)	raw_data.get('raw_index_html')	KirchenorgelForumCrawler.transform, line 365

8. Research use, citation and rights

No general permission to redistribute posts or attachments was established.[124]

S047. SONORA

Catalogue key: sonora | Role: registered source | Documentary unit: Archival item linked to organ, building, place or builder

1. Identity, scope and evidence

Source account. SONORA makes material from six Swedish archives accessible through map, gallery and archive views. Its stated selection follows Einar Erici's inventory and requires substantial original construction to survive; organs retaining only a facade are excluded.[125,126,127]

Attribution recorded in the catalogue: Göteborg International Organ Academy Association; University of Gothenburg / GRIDH. Scope: Swedish historical organs, primarily preserved instruments built before the 1860s.

2. Native record structure and contents

Selected document; 2026-09-11. JSON response linking an organ/work record to place, historical activities and archival documents[127]

- Place name and a distinct place identifier are returned.
- Historical maker and later intervention agents are represented.
- Construction, redisposition, restoration and inventory have separate activity records.
- Embedded HTML contains division stop names and pitches.
- Manual and pedal compasses are included in the specification fragment.
- Separate key and stop action fields are populated.
- Bellows type, quantity and descriptive text are populated.
- Console position and surviving facade-pipe information are present.
- Six archival document references include archive names and item identifiers.
- Historical activity records link to separate organ/work identifiers.

Repeated structures: Numbered activity objects and document objects accompany technical fields; specification is an embedded HTML fragment.

3. Discovery and acquisition

Place discovery and organ-detail retrieval are separate routes. Although housed in the static-crawler module, this subclass interprets JSON organ data together with cached place context and coordinates.[5]

Configured source address: https://sonora.dh.gu.se. This is a registry setting, not a new availability test.

Configured primary identifier: sonora_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- default_country: "Sweden" (aggregator/static_html_crawler.py, line 2089).

4. Extraction and normalization

The transform reads source organ fields, translates place labels into building/town context, parses builder/history and disposition material, and retains the organ response, place response and reference context together as raw evidence. Structured responses can still contain HTML specifications and documentary references.[5]

5. Dates, relationships and historical states

Historical activities use separate organ/work identifiers, dates and agents for one place.[127]

This survival criterion matters in comparison with inventories that include lost instruments. Archive, series and item references should accompany organ links. The portal's research-project link returned 404 even though the database interface remained accessible.[125,126,127]

6. Preservation, loss and uncertainty

The class name or module location does not determine the source format. Place IDs, organ/work IDs and documentary references must be distinguished. A joint raw-evidence object preserves the acquisition context but is an assembled representation, not necessarily a single source response.[5]

Limit of the selected-document observation: The JSON mixes record types and embeds HTML. A document count does not prove that document images were inspected.

7. Implementation evidence and rule register

Registered class: SonoraCrawler in aggregator/static_html_crawler.py; inheritance: SonoraCrawler > StaticHtmlOrganCrawler > BaseCrawler.[5]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- SonoraCrawler.crawl (lines 2091-2146).
- SonoraCrawler.transform (lines 2202-2275).
- SonoraCrawler._event_for_organ (lines 2345-2362).
- SonoraCrawler._events_from_organ_data (lines 2303-2316).
- SonoraCrawler._technical_rows (lines 2400-2412).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	sum((len(stops) for stops in stops_by_division.values()))	SonoraCrawler.transform, line 2242
manual_count	self._manual_count(stops_by_division)	SonoraCrawler.transform, line 2241
stops_by_division	stops_by_division	SonoraCrawler.transform, line 2240
raw_specification	raw_spec	SonoraCrawler.transform, line 2238
dates	[year] if year else []	SonoraCrawler.transform, line 2219
country	self.default_country	SonoraCrawler.transform, line 2227

8. Research use, citation and rights

A CC BY 4.0 notice accompanies a specific linked text; no single licence for all six archives was established.[125,126,127]

S048. Sonus Paradisi Free and Demo Sample Sets

Catalogue key: sonus_paradisi_free_stuff | Role: registered source | Documentary unit: Free or demonstration product announcement

1. Identity, scope and evidence

Source account. Aggregator targets the producer's Free Stuff blog category rather than the entire shop. The page sits within a catalogue organised by country and instrument, including full products outside the targeted category.[128]

Attribution recorded in the catalogue: Sonus Paradisi. Scope: Selected subset of a wider commercial organ catalogue.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects second-level headings and their associated content. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://www.sonusparadisi.cz/en/blog/category/free-stuff/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "free_demo_catalog" (aggregator/grandorgue_library_crawler.py, line 1335).

4. Extraction and normalization

The implemented unit is free or demonstration product sections. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Navigation links must not be counted as products in the selected collection. A demonstration's stop selection and release date belong to that edition, and should not be inherited from the corresponding full sample set.[128]

6. Preservation, loss and uncertainty

A demonstration may omit ranks, perspectives or other material from a commercial edition. Demo and full-product specifications cannot be interchanged. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: SonusParadisiFreeStuffCrawler in aggregator/grandorgue_library_crawler.py; inheritance: SonusParadisiFreeStuffCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- SonusParadisiFreeStuffCrawler.parse_html (lines 1337-1396).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	_first_int(_first_match(text, '(?i)(\\d+)\\s*stops?'))	SonusParadisiFreeStuffCrawler.parse_html, line 1383
manual_count	_first_int(_first_match(text, '(?i)(\\d+)\\s*manual'))	SonusParadisiFreeStuffCrawler.parse_html, line 1381
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'demo_sample_set' if is_demo else 'sample_set'	SonusParadisiFreeStuffCrawler.parse_html, line 1372
requires_sample_set	False	SonusParadisiFreeStuffCrawler.parse_html, line 1375
is_odf_only	False	SonusParadisiFreeStuffCrawler.parse_html, line 1374

8. Research use, citation and rights

The category name establishes neither a standard open licence nor unrestricted redistribution of samples.[128]

S049. SourceForge GrandOrgue Sample Sets

Catalogue key: sourceforge_ourorgan_samplesets | Role: registered source | Documentary unit: Historical sample-set directory entry

1. Identity, scope and evidence

Source account. The page explicitly says that it is no longer maintained and directs readers to the GrandOrgue GitHub wiki for the current list. It retains free/demo entries and definition files for commercial sets.[129]

Attribution recorded in the catalogue: GrandOrgue project contributors. Scope: Former GrandOrgue list on SourceForge.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects headings and nested list nodes in the Markdown content. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://sourceforge.net/p/ourorgan/samplesets/Sample%20Sets/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "legacy_registry" (aggregator/grandorgue_library_crawler.py, line 965).

4. Extraction and normalization

The implemented unit is items qualified by surrounding headings and parent-list context. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Keep this as a historical version and record its successor relationship. Counting it alongside the GitHub list as an independent current catalogue would duplicate much of the evidence and misstate maintenance status.[129]

6. Preservation, loss and uncertainty

Parent notes can qualify several files. A download directory is not a census of recorded physical instruments. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: SourceForgeOurOrganSamplesetsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: SourceForgeOurOrganSamplesetsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- SourceForgeOurOrganSamplesetsCrawler.parse_html (lines 967-1035).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	_first_int(_first_match(text, '(?i)(\\d+)\\s*stops?'))	SourceForgeOurOrganSamplesetsCrawler.parse_html, line 1022
manual_count	_first_int(_first_match(text, '(?i)(\\d+)\\s*manual'))	SourceForgeOurOrganSamplesetsCrawler.parse_html, line 1020
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'odf' if is_odf else 'download' if is_download else 'registry_entry'	SourceForgeOurOrganSamplesetsCrawler.parse_html, line 1011
requires_sample_set	is_odf	SourceForgeOurOrganSamplesetsCrawler.parse_html, line 1014
is_odf_only	is_odf	SourceForgeOurOrganSamplesetsCrawler.parse_html, line 1013

8. Research use, citation and rights

The page directs readers to individual agreements; the listed sample sets have no common licence established here.[129]

S050. Varhany

Catalogue key: varhany | Role: registered source | Documentary unit: Linked organ, builder, specification, photograph or literature record

1. Identity, scope and evidence

Source account. The author dates the private database to 1998 and its web publication to 2005. The about page describes literature, archival work and instrument inspection as inputs, with checked contributions rather than unrestricted direct editing. Some scanned literature is deliberately non-public.[130,131,132]

Attribution recorded in the catalogue: Štěpán Svoboda. Scope: Czech Republic.

2. Native record structure and contents

Selected document; 2026-09-11. Place-centred database page with event rows and expandable specifications[132]

- Place and church identify the record.
- Nearby organs are listed with distances and linked places.
- Builders appear in dated work rows.
- Repeated work rows contain dates, activities and occasional costs.
- A historical work row expands to a manual/pedal specification.
- Manual/pedal grouping and compact division/stop counts are supplied.
- Literature links are attached to individual work rows.
- Builders, literature and nearby places have separate links.
- A page-update timestamp is displayed.

Repeated structures: Each dated work row associates a builder, activity, specification/literature flags, cost and status.

Not established in this document check: Status codes occur in the work rows, but their meaning was not established here; no condition classification is assigned..

3. Discovery and acquisition

The crawler works with locality and organ references and retains both lok_id and varhany_id. HTML layout classes distinguish columns in the locality or editorial view; tooltips can supply abbreviated technical meanings.[133]

Configured source address: https://www.varhany.net. This is a registry setting, not a new availability test.

Configured primary identifier: varhany_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

Location, participants, specification, media, literature and links are extracted separately. The specification routine scans the type/work column for a Roman-manual-count and stop-count pattern, reads action information from the adjoining cell and tooltip, and stops at the first matching count row. If no count pattern is found, it retains a textual specification fallback.[133]

5. Dates, relationships and historical states

Many historical interventions occur within one place record, including predecessor and replacement instruments.[132]

Its separation of evidence sources is useful for tracing assertions. The site itself says that current-condition information is indicative and often based on literature. Such information needs its documentary date before use as a present-state observation.[130,131,132]

6. Preservation, loss and uncertainty

The first count-bearing row is not proof of the latest historical state. The inspected specification routine concentrates on summary counts and action context; a full source stoplist cannot be inferred from those counts. A locality identifier and an organ/work identifier should not be collapsed into one unit.[133]

Limit of the selected-document observation: Not every event has a stoplist. Event rows, organ identities and specification counts are different units.

7. Implementation evidence and rule register

Registered class: VarhanyCrawler in aggregator/varhany_crawler.py; inheritance: VarhanyCrawler > BaseCrawler.[133]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- VarhanyCrawler.crawl (lines 59-188).
- VarhanyCrawler.transform (lines 417-452).
- VarhanyCrawler._extract_specifications (lines 590-642).
- VarhanyCrawler._extract_literature (lines 680-693).
- VarhanyCrawler._extract_involved (lines 553-588).
- VarhanyCrawler._extract_links (lines 695-704).
- VarhanyCrawler._extract_locations (lines 478-551).
- VarhanyCrawler._extract_media (lines 644-678).
- VarhanyCrawler._extract_tooltip (lines 706-714).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
stop_count	match.group(2)	VarhanyCrawler._extract_specifications, line 608
manual_count	match.group(1)	VarhanyCrawler._extract_specifications, line 607
raw_str	bib	VarhanyCrawler._extract_literature, line 692
country (read)	raw_data.get('country')	VarhanyCrawler.transform, line 421
varhany_id (read)	raw_data.get('varhany_id')	VarhanyCrawler.transform, line 448
url (read)	raw_data.get('url')	VarhanyCrawler.transform, line 449

8. Research use, citation and rights

Copyright is asserted and some supporting scans are restricted; a general redistribution licence was not established.[130,131,132]

S051. VI-Control

Catalogue key: vi_control | Role: registered source | Documentary unit: Virtual-instrument discussion post or thread

1. Identity, scope and evidence

Source account. The introduction names Frederick Russ and describes a community for composers, producers and developers of virtual instruments. Organ discussions are selected from this much broader subject range.[134]

Attribution recorded in the catalogue: Frederick Russ / VI-Control. Scope: Broad composer and music-production community; organ material is a subset.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The adapter parses acquired forum-thread pages and supplies forum, thread and post context to a VMI-specific normalizer.[135,136]

Configured source address: https://vi-control.net. This is a registry setting, not a new availability test.

Configured primary identifier: thread_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

4. Extraction and normalization

Unlike the general forum-thread family, the transform iterates posts and returns a vmi_forum_thread_batch of individually normalized post records. Each carries author, links, media and thread context. Candidate product or entity mentions are downstream interpretations and must remain distinguishable from the post wording.[135,136]

5. Dates, relationships and historical states

Mentions of organ libraries can reveal terminology and user experience. They cannot establish product specifications without identifying the post, product version and relevant producer evidence. Global traffic statements are not counts of organ-related material.[134]

6. Preservation, loss and uncertainty

Post count, thread count, product mentions and actual products have different denominators. A recommendation or announcement does not validate a product specification. Quotations, publisher statements and other participants' comments require separate attribution.[135,136]

7. Implementation evidence and rule register

Registered class: VIControlForumCrawler in aggregator/vi_control_forum_crawler.py; inheritance: VIControlForumCrawler > BaseCrawler.[135,136]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- VIControlForumCrawler.crawl (lines 155-281).
- VIControlForumCrawler.transform (lines 347-383).
- VIControlForumCrawler._extract_links_and_media (lines 514-552).
- VIControlForumCrawler._extract_quotes (lines 555-571).
- VIControlForumCrawler._parse_forum_thread_rows (lines 442-471).
- VIControlForumCrawler._parse_post (lines 473-512).

- VIControlForumCrawler.parse_thread_html (lines 385-389).
- VIControlForumCrawler.parse_thread_pages (lines 391-440).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_forum_thread_batch'	VIControlForumCrawler.transform, line 379
source_snapshot	{'page_urls': [page.get('url') for page in pages if page.get('url')]}	VIControlForumCrawler.parse_thread_pages, line 439
html (read)	raw_data.get('html')	VIControlForumCrawler.transform, line 348
raw_html (read)	raw_data.get('raw_html')	VIControlForumCrawler.transform, line 348
url (read)	raw_data.get('url')	VIControlForumCrawler.transform, line 352
thread_url (read)	parsed['thread'].get('thread_url')	VIControlForumCrawler.transform, line 352

8. Research use, citation and rights

No general licence for redistributing discussions or participant data was established.[134]

S052. Virtual Pipe Organs Free Sample Sets and Producers

Catalogue key: virtual_pipe_organs_free_producers | Role: registered source | Documentary unit: Curated recommendation or producer link

1. Identity, scope and evidence

Source account. The page expressly describes its selection as non-exhaustive and directs readers to favourite sets, producers and other indexes. Some commentary refers to the producer landscape in early 2020.[137,138]

Attribution recorded in the catalogue: Virtual Pipe Organs. Scope: Selected free sample sets and discovery resources.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects list items within the entry-content area. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://virtualpipeorgans.wordpress.com/sample-sets/free-sample-set-producers/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "producer_registry" (aggregator/grandorgue_library_crawler.py, line 1100).

4. Extraction and normalization

The implemented unit is producer-directory entries. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

This is a dated curatorial selection. Its producer totals and availability statements should not be used as current market statistics. It overlaps with the GrandOrgue lists and with separately registered producer catalogues.[137,138]

6. Preservation, loss and uncertainty

A producer link is a discovery lead, not one sample set or one organ. Free access and redistribution rights are separate questions. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: VirtualPipeOrgansFreeProducersCrawler in aggregator/grandorgue_library_crawler.py; inheritance: VirtualPipeOrgansFreeProducersCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- VirtualPipeOrgansFreeProducersCrawler.parse_html (lines 1102-1145).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'odf_reference' if is_odf else 'producer_registry'	VirtualPipeOrgansFreeProducersCrawler.parse_html, line 1129
requires_sample_set	is_odf	VirtualPipeOrgansFreeProducersCrawler.parse_html, line 1132
is_odf_only	is_odf and 'sample set' not in lower	VirtualPipeOrgansFreeProducersCrawler.parse_html, line 1131

8. Research use, citation and rights

The page points to third-party products; permissions must be established at those products.[137,138]

S053. Virtual Pipe Organs ODFs For GrandOrgue

Catalogue key: virtual_pipe_organs_odfs | Role: registered source | Documentary unit: Organ definition file for external sample recordings

1. Identity, scope and evidence

Source account. The page explains that an ODF supplies the structure needed to play compatible recordings in GrandOrgue and links to adaptations for selected commercial sample sets. It distinguishes the definition file from the underlying recording material.[139,138]

Attribution recorded in the catalogue: Virtual Pipe Organs. Scope: GrandOrgue adaptations of selected plain-wave commercial sample sets.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

The source-specific parser selects links within the GrandOrgue ODF section. The library family can enrich discovered items from linked detail pages; catalogue and detail capture are distinct operations.[11,12]

Configured source address: https://virtualpipeorgans.wordpress.com/sample-sets/odfs-for-grandorgue/. This is a registry setting, not a new availability test.

Configured primary identifier: item_id. This names the adapter field; native and derived identifier semantics must be checked in the relevant rule.

Selected effective defaults (runtime settings can override these):

- catalog_kind: "odf_registry" (aggregator/grandorgue_library_crawler.py, line 1041).

4. Extraction and normalization

The implemented unit is definition-file catalogue entries. parse_html produces a collection of item records; the common transform wraps them in a vmi_library_batch. Source fragments, item identity, links and product context are passed to the VMI normalizer. The exact selectors and constructed fields are enumerated in the technical evidence register. The native wording remains necessary to distinguish a producer, a product, a version and the organ represented.[11,12]

Product names, software labels, access hints and organ-relation hints may include constants or text-based classifications supplied by the adapter. They are not automatically independent source assertions. The rule register distinguishes literals from values read out of a document.

5. Dates, relationships and historical states

Preserve the dependency on the exact external sample set. An ODF is not a complete audio library, and a downloadable ODF cannot be counted as an independently sampled organ.[139,138]

6. Preservation, loss and uncertainty

A definition file may require separately obtained samples; it is not by itself a complete recording library. The existence of a parser is not a fresh check of the live catalogue, and a parsed link is not evidence that its downloadable content was inspected.[11,12]

7. Implementation evidence and rule register

Registered class: VirtualPipeOrgansOdfsCrawler in aggregator/grandorgue_library_crawler.py; inheritance: VirtualPipeOrgansOdfsCrawler > GrandOrgueLibraryCrawler > BaseCrawler.[11,12]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- GrandOrgueLibraryCrawler.crawl (lines 86-136) [inherited].
- VirtualPipeOrgansOdfsCrawler.parse_html (lines 1043-1094).
- GrandOrgueLibraryCrawler.transform (lines 138-156) [inherited].

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
record_type	'vmi_library_batch'	GrandOrgueLibraryCrawler.transform, line 148
package_kind	'odf'	VirtualPipeOrgansOdfsCrawler.parse_html, line 1077
requires_sample_set	True	VirtualPipeOrgansOdfsCrawler.parse_html, line 1080
is_odf_only	True	VirtualPipeOrgansOdfsCrawler.parse_html, line 1079

8. Research use, citation and rights

The definition file and recordings require separate permissions; free access to one does not cover the other.[139,138]

S054. PDMX Public Domain MusicXML

Catalogue key: score_pdmx | Role: score collection | Documentary unit: Dataset row and associated symbolic or rendered files

1. Identity, scope and evidence

Source account. The pinned Zenodo record is version v9, published 1 June 2025. Its description reports 31,221 records with conflicting public and embedded licence metadata and recommends the no_license_conflict subset. It separately identifies files that could not be converted from corrupted originals.[140]

Attribution recorded in the catalogue: Phillip Long; Zachary Novack; Julian McAuley; Taylor Berg-Kirkpatrick. Scope: PDMX v9; broad music corpus with organ-relevant subsets.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

This is the pdmx collection route inside DigitalSheetMusicCrawler, not a separately registered crawler. The score-level route reads a supplied local dataset snapshot. A separate Zenodo-inventory route yields dataset/file metadata and an access-blocked status rather than automatically downloading the multi-gigabyte collection.[21,22]

4. Extraction and normalization

The shared score normalizer separates work metadata, assets, formats, collection context and licence statements. Dataset-row filtering and asset selection are separate functions. A dataset inventory item must not be counted as one successfully acquired score. Licence-conflict flags, failed conversions and duplicate representations require the dataset's own definitions.[21,22]

5. Dates, relationships and historical states

Retain the version DOI, row identifier, conflict flag and file format. An organ subset must be defined reproducibly; the full dataset size is not a count of organ scores. The authors request citations to both their paper and their collaborators' paper when the dataset is used.[140]

6. Preservation, loss and uncertainty

Retain the version DOI, row identifier, conflict flag and file format. An organ subset must be defined reproducibly; the full dataset size is not a count of organ scores. The authors request citations to both their paper and their collaborators' paper when the dataset is used. Availability of a collection route does not demonstrate that all pages or files were acquired.[21,22]

7. Implementation evidence and rule register

Registered class: DigitalSheetMusicCrawler in aggregator/digital_sheet_music_crawler.py; inheritance: DigitalSheetMusicCrawler > BaseCrawler.[21,22]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- DigitalSheetMusicCrawler.crawl (lines 145-164).
- DigitalSheetMusicCrawler.transform (lines 166-179).
- DigitalSheetMusicCrawler._extract_tar_gz_member_from_url (lines 780-812).
- DigitalSheetMusicCrawler._parse_mutopia_listing (lines 907-988).
- DigitalSheetMusicCrawler._respectful_pause (lines 1021-1024).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
score_id (read)	raw_data.get('score_id')	DigitalSheetMusicCrawler.transform, line 167
id (read)	raw_data.get('id')	DigitalSheetMusicCrawler.transform, line 167
raw_data_json (read)	payload.get('raw_data_json')	DigitalSheetMusicCrawler.transform, line 169
transform_run_id (read)	raw_data.get('transform_run_id')	DigitalSheetMusicCrawler.transform, line 178
task_id (read)	kwargs.get('task_id')	DigitalSheetMusicCrawler.crawl, line 148
ok (read)	preflight.get('ok')	DigitalSheetMusicCrawler.crawl, line 150

8. Research use, citation and rights

The record carries CC BY 4.0 and reports item-level licence conflicts; its title is not proof that every file is free of conflicting claims.[140]

S055. Little Organ Book / Orgelbuechlein Dataset

Catalogue key: score_orgelbuechlein | Role: score collection | Documentary unit: Encoded score and harmonic-analysis version

1. Identity, scope and evidence

Source account. The repository provides MusicXML, PDF and Roman-numeral-text material. It identifies Pierre Gouin's 2008 edition as the score source and credits the 2022 conversion, editing and annotation. Its harmonic analysis is presented as an interpretation, and a revised BWV 621 text analysis coexists with the earlier version.[141]

Attribution recorded in the catalogue: H. Asai; Y. Gao; S. Shimosaka; S. Tojo; Y. Uehara. Scope: Bach's Orgelbüchlein.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

This is the orgelbuechlein collection route inside DigitalSheetMusicCrawler, not a separately registered crawler. The route lists repository files and selects MusicXML extensions. The inspected code assigns Bach as composer, organ instrumentation and a licence label, and constructs a corresponding PDF URL by replacing a directory and suffix.[21,22]

4. Extraction and normalization

The shared score normalizer separates work metadata, assets, formats, collection context and licence statements. Those constants and the constructed PDF address are implementation values. Their correctness needs the collection documentation and an actual asset check; the existence of a generated URL does not prove a file exists. Score edition, encoding and analytical annotation remain separate.[21,22]

5. Dates, relationships and historical states

Cite the edition, encoding and analysis separately. Corrected notation and harmonic labels are editorial contributions rather than a direct transcription of compositional intent. Pin a commit and file path for repeatable analysis.[141]

6. Preservation, loss and uncertainty

Cite the edition, encoding and analysis separately. Corrected notation and harmonic labels are editorial contributions rather than a direct transcription of compositional intent. Pin a commit and file path for repeatable analysis. Availability of a collection route does not demonstrate that all pages or files were acquired.[21,22]

7. Implementation evidence and rule register

Registered class: DigitalSheetMusicCrawler in aggregator/digital_sheet_music_crawler.py; inheritance: DigitalSheetMusicCrawler > BaseCrawler.[21,22]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- DigitalSheetMusicCrawler.crawl (lines 145-164).
- DigitalSheetMusicCrawler.transform (lines 166-179).
- DigitalSheetMusicCrawler._extract_tar_gz_member_from_url (lines 780-812).
- DigitalSheetMusicCrawler._parse_mutopia_listing (lines 907-988).
- DigitalSheetMusicCrawler._respectful_pause (lines 1021-1024).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
score_id (read)	raw_data.get('score_id')	DigitalSheetMusicCrawler.transform, line 167
id (read)	raw_data.get('id')	DigitalSheetMusicCrawler.transform, line 167
raw_data_json (read)	payload.get('raw_data_json')	DigitalSheetMusicCrawler.transform, line 169
transform_run_id (read)	raw_data.get('transform_run_id')	DigitalSheetMusicCrawler.transform, line 178
task_id (read)	kwargs.get('task_id')	DigitalSheetMusicCrawler.crawl, line 148
ok (read)	preflight.get('ok')	DigitalSheetMusicCrawler.crawl, line 150

8. Research use, citation and rights

The README states CC BY-NC 3.0 and identifies the original edition and adaptation credits.[141]

S056. IMSLP / Werner Icking Music Collection

Catalogue key: score_imslp_wima | Role: score collection | Documentary unit: Edition or file attached to a work page

1. Identity, scope and evidence

Source account. Aggregator groups IMSLP and WIMA under one collection key and uses selected work pages. The MusicXML resource directory notes that the Werner Icking archive merged with IMSLP while its original site remains available.[142,19,20]

Attribution recorded in the catalogue: IMSLP / Petrucci Music Library; Werner Icking Music Archive contributors. Scope: Broad score collections; organ-relevant selections.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

This is the imslp_wima collection route inside DigitalSheetMusicCrawler, not a separately registered crawler. The collection route handles selected work-page material and uses the shared HTML asset finder. It is a route to selected pages, not a complete mirror of IMSLP or the historical WIMA site.[21,22]

4. Extraction and normalization

The shared score normalizer separates work metadata, assets, formats, collection context and licence statements. A work page may contain multiple editions, arrangements and file types. The collection key is not an umbrella rights statement or a stable edition identifier.[21,22]

5. Dates, relationships and historical states

A work page can host several editions, arrangements and file formats. Cite the specific file, editor or contributor and edition date; a work title alone is insufficient. WIMA provenance should survive the hosting relationship.[142,19,20]

6. Preservation, loss and uncertainty

A work page can host several editions, arrangements and file formats. Cite the specific file, editor or contributor and edition date; a work title alone is insufficient. WIMA provenance should survive the hosting relationship. Availability of a collection route does not demonstrate that all pages or files were acquired.[21,22]

7. Implementation evidence and rule register

Registered class: DigitalSheetMusicCrawler in aggregator/digital_sheet_music_crawler.py; inheritance: DigitalSheetMusicCrawler > BaseCrawler.[21,22]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- DigitalSheetMusicCrawler.crawl (lines 145-164).
- DigitalSheetMusicCrawler.transform (lines 166-179).
- DigitalSheetMusicCrawler._extract_tar_gz_member_from_url (lines 780-812).
- DigitalSheetMusicCrawler._parse_mutopia_listing (lines 907-988).
- DigitalSheetMusicCrawler._respectful_pause (lines 1021-1024).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
score_id (read)	raw_data.get('score_id')	DigitalSheetMusicCrawler.transform, line 167
id (read)	raw_data.get('id')	DigitalSheetMusicCrawler.transform, line 167
raw_data_json (read)	payload.get('raw_data_json')	DigitalSheetMusicCrawler.transform, line 169
transform_run_id (read)	raw_data.get('transform_run_id')	DigitalSheetMusicCrawler.transform, line 178
task_id (read)	kwargs.get('task_id')	DigitalSheetMusicCrawler.crawl, line 148
ok (read)	preflight.get('ok')	DigitalSheetMusicCrawler.crawl, line 150

8. Research use, citation and rights

Rights must be read for the specific edition, file and applicable jurisdiction. This review does not assign a collection-wide licence.[142,19,20]

S057. Mutopia Project

Catalogue key: score_mutopia | Role: score collection | Documentary unit: Typeset edition with editable and rendered files

1. Identity, scope and evidence

Source account. Mutopia supplies PDF, MIDI and editable LilyPond files and credits the typesetter of each piece. Its licence page distinguishes public-domain contributions from several versions of Creative Commons Attribution and Attribution-ShareAlike licences.[143,144]

Attribution recorded in the catalogue: The Mutopia Project and individual typesetters. Scope: General music collection containing organ works.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

This is the mutopia collection route inside DigitalSheetMusicCrawler, not a separately registered crawler. The source-specific listing parser reads tabular title, composer and catalogue-number cells, an instrumentation row, piece-information links and labelled assets. A separate continuation helper handles following listing pages.[21,22]

4. Extraction and normalization

The shared score normalizer separates work metadata, assets, formats, collection context and licence statements. The code recognises specific licence-link labels and asset labels. Unrecognised labels need review, and pagination support does not prove that a crawl reached every page. The editable LilyPond source, PDF and MIDI are different representations of a contribution.[21,22]

5. Dates, relationships and historical states

The edition and typesetter belong in the citation alongside the composer. The home-page music count is not an organ-work count. Keep the source file and generated performance or print format distinct.[143,144]

6. Preservation, loss and uncertainty

The edition and typesetter belong in the citation alongside the composer. The home-page music count is not an organ-work count. Keep the source file and generated performance or print format distinct. Availability of a collection route does not demonstrate that all pages or files were acquired.[21,22]

7. Implementation evidence and rule register

Registered class: DigitalSheetMusicCrawler in aggregator/digital_sheet_music_crawler.py; inheritance: DigitalSheetMusicCrawler > BaseCrawler.[21,22]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- DigitalSheetMusicCrawler.crawl (lines 145-164).
- DigitalSheetMusicCrawler.transform (lines 166-179).
- DigitalSheetMusicCrawler._extract_tar_gz_member_from_url (lines 780-812).
- DigitalSheetMusicCrawler._parse_mutopia_listing (lines 907-988).
- DigitalSheetMusicCrawler._respectful_pause (lines 1021-1024).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
score_id (read)	raw_data.get('score_id')	DigitalSheetMusicCrawler.transform, line 167
id (read)	raw_data.get('id')	DigitalSheetMusicCrawler.transform, line 167
raw_data_json (read)	payload.get('raw_data_json')	DigitalSheetMusicCrawler.transform, line 169
transform_run_id (read)	raw_data.get('transform_run_id')	DigitalSheetMusicCrawler.transform, line 178
task_id (read)	kwargs.get('task_id')	DigitalSheetMusicCrawler.crawl, line 148
ok (read)	preflight.get('ok')	DigitalSheetMusicCrawler.crawl, line 150

8. Research use, citation and rights

Multiple licences apply by piece; the specific edition determines the obligations.[143,144]

S058. MuseScore.com Direct

Catalogue key: score_musescore_direct | Role: score collection | Documentary unit: User-published score version

1. Identity, scope and evidence

Source account. The local collection definition requires credentials and accepts explicit MuseScore URLs. The public entry-point check encountered an access challenge; no score content or current download terms were verified.[145,20]

Attribution recorded in the catalogue: MuseScore.com and individual score contributors. Scope: Directly selected score URLs.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

This is the musescore_direct collection route inside DigitalSheetMusicCrawler, not a separately registered crawler. The collection definition marks the direct route as requiring credentials and the implementation accepts explicitly selected score addresses.[21,22]

4. Extraction and normalization

The shared score normalizer separates work metadata, assets, formats, collection context and licence statements. The dated source check encountered an access challenge. No source-content inspection, current download entitlement or successful acquisition is established by the configured route.[21,22]

5. Dates, relationships and historical states

Record contributor, score identifier, revision and downloaded format for any later acquisition. Platform identity does not identify a score's editor, and a public-domain composition does not determine the rights in its modern encoding.[145,20]

6. Preservation, loss and uncertainty

Record contributor, score identifier, revision and downloaded format for any later acquisition. Platform identity does not identify a score's editor, and a public-domain composition does not determine the rights in its modern encoding. Availability of a collection route does not demonstrate that all pages or files were acquired.[21,22]

7. Implementation evidence and rule register

Registered class: DigitalSheetMusicCrawler in aggregator/digital_sheet_music_crawler.py; inheritance: DigitalSheetMusicCrawler > BaseCrawler.[21,22]

Entry points and source-specific routines (full eligible inventory in the JSON companion):

- DigitalSheetMusicCrawler.crawl (lines 145-164).
- DigitalSheetMusicCrawler.transform (lines 166-179).
- DigitalSheetMusicCrawler._extract_tar_gz_member_from_url (lines 780-812).
- DigitalSheetMusicCrawler._parse_mutopia_listing (lines 907-988).
- DigitalSheetMusicCrawler._respectful_pause (lines 1021-1024).

These are code constructions or reads, not a verified source-to-final-schema crosswalk.

Key / role	Inspected expression	Code locator
score_id (read)	raw_data.get('score_id')	DigitalSheetMusicCrawler.transform, line 167
id (read)	raw_data.get('id')	DigitalSheetMusicCrawler.transform, line 167
raw_data_json (read)	payload.get('raw_data_json')	DigitalSheetMusicCrawler.transform, line 169
transform_run_id (read)	raw_data.get('transform_run_id')	DigitalSheetMusicCrawler.transform, line 178
task_id (read)	kwargs.get('task_id')	DigitalSheetMusicCrawler.crawl, line 148
ok (read)	preflight.get('ok')	DigitalSheetMusicCrawler.crawl, line 150

8. Research use, citation and rights

File-specific permissions and current account conditions remain unverified.[145,20]

S059. ORKASA

Catalogue key: audit_orkasa | Role: audit only | Documentary unit: Church organ database; no registered Aggregator adapter

1. Identity, scope and evidence

Source account. The church's organ-information page identifies ORKASA as an organ database. Aggregator's dated investigation classified its access as requiring an external application client rather than ordinary HTTP acquisition.[146,147]

Attribution recorded in the catalogue: Evangelisch-Lutherische Landeskirche Sachsens. Scope: Organs associated with the Saxon regional church.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

No independently registered crawler is present for this audit-only entry. The cited investigation and public source account define the available evidence.

4. Extraction and normalization

No source-specific transformation is attributed to this resource. Inferring a field model from the common MODAVIS schema would manufacture evidence.

5. Dates, relationships and historical states

That classification is a recorded implementation boundary, not a judgement of research value. ORKASA and the Dresden-led Saxony inventory have different institutional scopes and must remain separate entries.[146,147]

6. Preservation, loss and uncertainty

That classification is a recorded implementation boundary, not a judgement of research value. ORKASA and the Dresden-led Saxony inventory have different institutional scopes and must remain separate entries.

7. Implementation evidence and rule register

Not applicable: no registered adapter. No inferred fields, selectors, mappings or test results are supplied.

8. Research use, citation and rights

No data-access or redistribution permission was established in this review.[146,147]

S060. Encyclopedia of Organ Stops

Catalogue key: audit_organstops | Role: audit only | Documentary unit: Organ-stop concept entry

1. Identity, scope and evidence

Source account. The encyclopedia provides alphabetical access to stop terminology and names Edward L. Stauff as owner, operator and copyright holder. Aggregator's investigation excluded it from the organ-instance acquisition lane.[148,147]

Attribution recorded in the catalogue: Edward L. Stauff. Scope: Encyclopedia of Organ Stops.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

No independently registered crawler is present for this audit-only entry. The cited investigation and public source account define the available evidence.

4. Extraction and normalization

No source-specific transformation is attributed to this resource. Inferring a field model from the common MODAVIS schema would manufacture evidence.

5. Dates, relationships and historical states

That exclusion concerns record type. The encyclopedia remains relevant to terminology research, but a stop-name definition cannot identify a physical organ or prove that two historical stop names denote the same construction.[148,147]

6. Preservation, loss and uncertainty

That exclusion concerns record type. The encyclopedia remains relevant to terminology research, but a stop-name definition cannot identify a physical organ or prove that two historical stop names denote the same construction.

7. Implementation evidence and rule register

Not applicable: no registered adapter. No inferred fields, selectors, mappings or test results are supplied.

8. Research use, citation and rights

The site asserts copyright; no general open licence was established.[148,147]

S061. Welt-Orgel-Archiv

Catalogue key: audit_welt_orgel_archiv | Role: audit only | Documentary unit: Unavailable historical web resource

1. Identity, scope and evidence

Source account. The endpoint recorded in Aggregator's investigation returned HTTP 404 again on 10 September 2026. No content was recovered and no registered adapter is present.[149,147]

Attribution recorded in the catalogue: Welt-Orgel-Archiv (title recorded in the investigation audit). Scope: Not recoverable from the checked endpoint.

2. Native record structure and contents

No separately documented detail-page observation is supplied for this entry. Its source account above describes the consulted portal or collection material. The code reading below is evidence about the adapter, not a substitute for inspecting an upstream record.

3. Discovery and acquisition

No independently registered crawler is present for this audit-only entry. The cited investigation and public source account define the available evidence.

4. Extraction and normalization

No source-specific transformation is attributed to this resource. Inferring a field model from the common MODAVIS schema would manufacture evidence.

5. Dates, relationships and historical states

Retain the URL and dated failure as evidence of the investigation's scope. The response does not establish the collection's historical extent, ownership, permanent disappearance or availability at another address.[149,147]

6. Preservation, loss and uncertainty

Retain the URL and dated failure as evidence of the investigation's scope. The response does not establish the collection's historical extent, ownership, permanent disappearance or availability at another address.

7. Implementation evidence and rule register

Not applicable: no registered adapter. No inferred fields, selectors, mappings or test results are supplied.

8. Research use, citation and rights

Not assessable from the error response.[149,147]

Shared helpers and interpretation boundaries

The following modules supplement the registered class inventory. They are indexed independently in the JSON companion, so a helper's supported fields are not falsely assigned to every source. A later offline recovery helper is not automatically part of an earlier frozen transfer.

- aggregator/transfer_schema.py: 31 function definitions indexed.[4]
- aggregator/digital_sheet_music_transfer_schema.py: 12 function definitions indexed.[22]
- aggregator/vmi_library_transfer_schema.py: 17 function definitions indexed.[12]
- aggregator/tuning_system_transfer_schema.py: 13 function definitions indexed.[16]
- aggregator/gdo_literature_transfer_schema.py: 6 function definitions indexed.[30]
- aggregator/institution_directory_transfer_schema.py: 7 function definitions indexed.[33]
- aggregator/ridim_iconography_transfer_schema.py: 14 function definitions indexed.[123]
- aggregator/forum_transfer_schema.py: 17 function definitions indexed.[45]
- aggregator/vmi_forum_transfer_schema.py: 32 function definitions indexed.[136]
- aggregator/musixplora/graph_builder.py: 36 function definitions indexed.[55]
- aggregator/musixplora/transfer_schema.py: 10 function definitions indexed.[56]
- aggregator/musixplora/organ_projection.py: 1 function definitions indexed.[57]
- aggregator/ohs_stoplist_text.py: 12 function definitions indexed.[72]
- aggregator/ohs_event_participants.py: 2 function definitions indexed.[73]
- aggregator/schmidt_history.py: 4 function definitions indexed.[88]
- aggregator/indo_event_participants.py: 1 function definitions indexed.[150]
- aggregator/swiss_specification_accounts.py: 11 function definitions indexed.[151]
- aggregator/swiss_configuration_date.py: 6 function definitions indexed.[152]

References

1. MODAVIS project. MODAVIS Aggregator: source-selection frame. Registry at revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; aggregator/__init__.py, CRAWLERS dictionary; SHA-256 77a20a2ecb41b01176edab8a978ea684223429fb69c09c740390696448c7701c. Exact registered keys and collection roles are preserved in evidence.json. Access to the underlying private repository is not required to inspect this selection frame.

2. Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, and Kate Crawford. Datasheets for Datasets. arXiv:1803.09010v8, revised 1 December 2021. https://arxiv.org/abs/1803.09010v8. Accessed 11 September 2026.

3. Groth, Paul, and Luc Moreau, eds. An Overview of the PROV Family of Documents. W3C Working Group Note, 30 April 2013. https://www.w3.org/TR/2013/NOTE-prov-overview-20130430/. Accessed 11 September 2026.

4. MODAVIS Aggregator. aggregator/transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 66e7c54c824cf418966ca7c8542e408217cf80130f1e447b7395f91e7b9d31d2. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

5. MODAVIS Aggregator. aggregator/static_html_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 f47fb9679d745a7731ee53a7983377a37b0092f7dea66a3500e83cd5181419aa. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/static_html_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

6. Steve Layden (website creator); Jeff Scofield (website manager). Aeolian-Skinner Archives. https://archive.pipeorgandatabase.org/aeolianskinner/main.html. Consulted 2026-09-10. bounded_direct_http_get

7. Steve Layden (website creator); Jeff Scofield (website manager). Opus 511: First Baptist Church (Murfreesboro, TN). https://archive.pipeorgandatabase.org/aeolianskinner/Specs/Op00511.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

8. Pottjewijd, Geert Jan. Homepage Organs of Arp Schnitger and organs in Drenthe. https://arpschnitger.nl/. Consulted 2026-09-10. bounded_direct_http_get

9. Geert Jan Pottjewijd. Former organ Glückstadt Stadtkirche. https://arpschnitger.nl/sgluckst.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

10. Augustine's Virtual Organs. List of AVO Hauptwerk sound samples. https://hauptwerk-augustine.info/Sample_sets.php. Consulted 2026-09-10. bounded_direct_http_get

11. MODAVIS Aggregator. aggregator/grandorgue_library_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 8654fc0b5874b2b520ce43cef5c43fd65c03c97307b1b39fbc1acca68adc4b65. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/grandorgue_library_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

12. MODAVIS Aggregator. aggregator/vmi_library_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 752fb6c6e0a6991cfeb12115fb3b2a20ebe1e8d0be9f191f782004c36e971d0e. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/vmi_library_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

13. Hans Eugen Frischknecht; Jakob Schmid; hosted by Bund Deutscher Orgelbaumeister. Stimmungen. https://bund-deutscher-orgelbaumeister.de/stimmungen/. Consulted 2026-09-10. bounded_direct_http_get

14. Frischknecht, Hans Eugen; Schmid, Jakob. Stimmungen Autoren. https://bund-deutscher-orgelbaumeister.de/stimmungen/authors.htm. Consulted 2026-09-10. bounded_direct_http_get

15. MODAVIS Aggregator. aggregator/bdo_tuning_systems_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 0671e570252fa9c8722bfc606add24bb5b7e41ddd44da821e4fad49be51a1b1f. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/bdo_tuning_systems_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

16. MODAVIS Aggregator. aggregator/tuning_system_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 29cb7c0b01f7d2311be94ca91ff48e78b7df2b1375226bf4c1cd180655a9ad42. Repository

locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/tuning_system_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

17. Dietrich Wierczeyko / Binauralpipes. Instrumente - Binauralpipes.com. https://binauralpipes.com/instrumente. Consulted 2026-09-10. bounded_direct_http_get

18. Contrebombarde Concert Hall. contrebombarde.com concert hall. https://www.contrebombarde.com/concerthall/organs/search. Consulted 2026-09-10. bounded_direct_http_get

19. MakeMusic, Inc.. MusicXML Sites: A List of Sites with Sheet Music in MusicXML Format. https://www.musicxml.com/music-in-musicxml/. Consulted 2026-09-10. bounded_direct_http_get

20. MODAVIS project. MODAVIS Aggregator: SOURCE_COLLECTIONS. Version 2bb6e85282e3ab11af1bcc7126290f67af176f53. Packaged extract in evidence.json; source path aggregator/digital_sheet_music_transfer_schema.py; SHA-256 dd652091e15347c1973c7b71613d7db0c5658db9307c9bfaf87c4cf7ed7de04e

21. MODAVIS Aggregator. aggregator/digital_sheet_music_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 4e8ac26682e90c7d14fce8f1d69edda9beb571f8c8bc7adce368d88183330694. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/digital_sheet_music_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

22. MODAVIS Aggregator. aggregator/digital_sheet_music_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 dd652091e15347c1973c7b71613d7db0c5658db9307c9bfaf87c4cf7ed7de04e. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/digital_sheet_music_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

23. Matrix Associates / The Estey Organ virtual museum. Estey Pipe Organs Brattleboro Vermont Virtual Museum. https://archive.pipeorgandatabase.org/estey/. Consulted 2026-09-10. bounded_direct_http_get

24. Matrix Associates / The Estey Organ virtual museum. Opus 1 - First Methodist Church, Brattleboro, VT. https://archive.pipeorgandatabase.org/estey/Opus1.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

25. Evensong LLC. Samplesets for Hauptwerk - EVENSONG LLC. https://www.evensongmusic.net/?page_id=155. Consulted 2026-09-10. bounded_direct_http_get

26. Evensong LLC. LICENSE AGREEMENT - HW SAMPLES - EVENSONG LLC. https://www.evensongmusic.net/?page_id=3884. Consulted 2026-09-10. bounded_direct_http_get

27. Reichling, Alfred; Reichling, Matthias. Literaturdatenbank - Gesellschaft der Orgelfreunde. https://www.gdo.de/recherchen/literaturdatenbank. Consulted 2026-09-10. bounded_direct_http_get

28. Alfred Reichling; Matthias Reichling; Gesellschaft der Orgelfreunde. GDO Literaturdatenbank (entry-point check). https://literaturdb.gdo.de/bib.php. Consulted 2026-09-10. bounded_direct_http_get; HTTP 500; page content not verified

29. MODAVIS Aggregator. aggregator/gdo_literature_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 0acff63fc05fb9ebb8e24b440755431d13ecc3bc24b800875c1db0f6b30856fa. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/gdo_literature_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

30. MODAVIS Aggregator. aggregator/gdo_literature_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 181dce7bff83d952eb4036fa00962f53e61be4f31c5148d79f60c25252027b10. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/gdo_literature_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

31. Gesellschaft der Orgelfreunde. Orgelbauer - Gesellschaft der Orgelfreunde. https://www.gdo.de/recherchen/links/orgelbauer. Consulted 2026-09-10. bounded_direct_http_get

32. MODAVIS Aggregator. aggregator/gdo_orgelbauer_links_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 30f4304bde50591faeaad6b6beb79421129cd4a26bef4800a299dc1810300c5d. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/gdo_orgelbauer_links_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

33. MODAVIS Aggregator. aggregator/institution_directory_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 a069a01be1aab0f7bd196efe85200f900c11add052f5aee2e7fd8674d8163aeb. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/institution_directory_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

34. GrandOrgue project contributors. Sample Sets · GrandOrgue/grandorgue Wiki · GitHub. https://github.com/GrandOrgue/grandorgue/wiki/Sample-Sets. Consulted 2026-09-10. bounded_direct_http_get

35. Orgue en France. Inventaire national des orgues. https://inventaire-des-orgues.fr/. Consulted 2026-09-10. bounded_direct_http_get

36. Orgue en France. Orgue à rouleaux église Saint-Pierre - Airvault, Deux-Sèvres. https://inventaire-des-orgues.fr/detail/orgue-airvault-eglise-saint-pierre-fr-79005-airva-stpier1-x/. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

37. MODAVIS Aggregator. aggregator/indo_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 dd26703917e4a6ed86e523c05596b40727d5df9ec55e88bcf46b4430283b01d0. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/indo_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

38. Instituto de Órganos Históricos de Oaxaca. IOHIO. https://www.iohio.org.mx/. Consulted 2026-09-10. bounded_direct_http_get

39. Instituto de Órganos Históricos de Oaxaca. Home | Explore, Conserve, Enjoy Now - IOHIO. https://www.iohio.org.mx/en/home. Consulted 2026-09-10. bounded_direct_http_get

40. Instituto de Órganos Históricos de Oaxaca. Basílica de la Soledad - IOHIO. https://www.iohio.org.mx/en/basilica-la-soledad. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

41. MODAVIS Aggregator. aggregator/iohio_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 56703e428800e98ec934346b7307b484e3293e584990412a241dbb3ad5f2a5e8. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/iohio_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

42. Jeux d'orgues. Jeux d'orgues. https://www.jeuxdorgues.com/. Consulted 2026-09-10. bounded_direct_http_get

43. Kirchenorgel-Forum. Kirchenorgel-Forum. https://www.kirchenorgelforum.at/. Consulted 2026-09-10. bounded_direct_http_get

44. MODAVIS Aggregator. aggregator/kirchenorgelforum_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 d008e9c0d121af88c1dbc5a710a013e2e2125f08f08e1670b4eb31dfa49b5469. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/kirchenorgelforum_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

45. MODAVIS Aggregator. aggregator/forum_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 19bbe595bd263c5393dbceb9c3c369e0f9a26ede9a65f09934b0213cd4fa3de8. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/forum_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

46. Palo, Lars. My free GrandOrgue samplesets - Lars virtual pipe organ site. https://familjenpalo.se/vpo/download/. Consulted 2026-09-10. bounded_direct_http_get

47. Lavender Audio. Lavender Audio - Hereford Cathedral Sample Sets. https://www.lavenderaudio.co.uk/organs/hereford/hwss.html. Consulted 2026-09-10. bounded_direct_http_get

48. Milan Digital Audio. Instruments: Milan Digital Audio. https://milandigitalaudio.com/instruments/. Consulted 2026-09-10. search_index_of_official_page; Instrument catalogue product headings; indexed age reported as last month

49. Milan Digital Audio. Milan Digital Audio Instruments (entry-point check). https://milandigitalaudio.com/instruments/. Consulted 2026-09-10. bounded_direct_http_get; HTTP 403; page content not verified

50. MPS Orgelseite contributors. Filebase - MPS Orgelseite. https://mps-orgelseite.de/home/filebase/. Consulted 2026-09-10. bounded_direct_http_get

51. Musique et musiciens / Music and musicians. Musique et musiciens / Music and musicians. http://www.musiqueorguequebec.ca/. Consulted 2026-09-10. bounded_direct_http_get

52. Musique et musiciens / Music and musicians. Église Sainte-Cécile, Le Bic, (QC). http://www.musiqueorguequebec.ca/orgues/quebec/bicsc.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

53. musiXplora / Universität Leipzig. musiXplora: Über den musiXplora. https://musixplora.de/about/. Consulted 2026-09-10. rendered_browser_page; About text, domain navigation; person-search interface at /musici/; imprint links at /impressum/

54. MODAVIS Aggregator. aggregator/musixplora/crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 4a04e0e7e33fba162e6920cc4aa1a2be4fbde24738b40f5982fe646a6a3afeed. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/musixplora/crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

55. MODAVIS Aggregator. aggregator/musixplora/graph_builder.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 9f4c4bbb280a35561c48ae57cb5e445840bb22a7056372ac32b6a6ebd2a93de2. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/musixplora/graph_builder.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

56. MODAVIS Aggregator. aggregator/musixplora/transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 634d2902308582cbfdfb40d8dd375652c36b4d14a01e8b7f523c4276313bba03. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/musixplora/transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

57. MODAVIS Aggregator. aggregator/musixplora/organ_projection.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 25623ec347a56ffb98f8335498f1b772e96b87bab469c3f4789de7bcc2410eed. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/musixplora/organ_projection.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

58. Norddeutsche Orgelmusikkultur in Niedersachsen und Europa e.V. (NOMINE). Historische Orgeln in Norddeutschland entdecken - NOMINE e.V.. https://nomine.net/. Consulted 2026-09-10. bounded_direct_http_get

59. Norddeutsche Orgelmusikkultur in Niedersachsen und Europa e.V. (NOMINE). Abbehausen, St. Laurentius - Nomine. https://nomine.net/orgel/abbehausen-st-laurentius/. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

60. MODAVIS Aggregator. aggregator/nomine_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 db87891d26884e5bf316dfc8051046ee0f37a2268dfbf2706d5226e8e10953e0. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/nomine_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

61. British Institute of Organ Studies. The NPOR | Home. https://npor.org.uk/. Consulted 2026-09-10. rendered_browser_page; Home page: provider heading, resource cards and footer

62. British Institute of Organ Studies (BIOS). NPOR - selected record (descriptive title). https://npor.org.uk/survey/D00002. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

63. MODAVIS Aggregator. aggregator/npor_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 37fed2b419954637942f223c61fbea73a3259acc101ba69f7293d6c734b2b525. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/npor_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

64. New York City Chapter, American Guild of Organists. The New York City Organ Project: Organs in the Borough of Manhattan. https://www.nycago.org/Organs/NYC/. Consulted 2026-09-10. search_index_of_official_page; Page introduction and borough navigation; indexed age reported as 2.9 years

65. New York City Chapter, American Guild of Organists. nycago_catalogue. https://www.nycago.org/Organs/NYC/. Consulted 2026-09-10. bounded_direct_http_get; HTTP 403; page content not verified

66. New York City Chapter, American Guild of Organists. NYC AGO - selected document (descriptive title). https://nycago.org/Organs/Brx/html/FordhamSkatingRink.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

67. MODAVIS Aggregator. aggregator/nycago_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 01c3d3ab2673ddff38c2c0f6b52ea347cca0f53dedca1183bf6cd510e1a9e16d. Repository locator: https://git

hub.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/nycago_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

68. Organ Historical Society. Pipe Organ Database: Explore Historic Pipe Organs. https://pipeorgandatabase.org/. Consulted 2026-09-10. web_reader_of_official_page; Home page counters and project footer

69. Organ Historical Society. Pipe Organ Database | Explore Historic Pipe Organs. https://pipeorgandatabase.org/. Consulted 2026-09-10. bounded_direct_http_get

70. Organ Historical Society. OHS Pipe Organ Database - selected record (descriptive title). https://pipeorgandatabase.org/instruments/1. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

71. MODAVIS Aggregator. aggregator/ohs_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 b38ae2df15a04d7dd7e74b88443223c50eca797c68b802cdc7ed3da96314f81b. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/ohs_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

72. MODAVIS Aggregator. aggregator/ohs_stoplist_text.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 8e4a28f9e605e7912207d26b587afc96b14a68ea48c5068fbd122f62366f1ae3. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/ohs_stoplist_text.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

73. MODAVIS Aggregator. aggregator/ohs_event_participants.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 af6da10cbf129e1fa0819132241fdc19e6428e48a3e0cd59faf60beeda737dce. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/ohs_event_participants.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

74. Organ Historical Trust of Australia. Organ Historical Trust of Australia. https://ohta.org.au/. Consulted 2026-09-10. bounded_direct_http_get

75. Organ Historical Trust of Australia. Copyright and privacy policy - Organ Historical Trust of Australia. https://ohta.org.au/copyright-privacy-policy/. Consulted 2026-09-10. bounded_direct_http_get

76. Organ Historical Trust of Australia. St Andrew's Lutheran Church Wickham Terrace. https://ohta.org.au/organs/organs/BrisbaneLutheran.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

77. Organ Index contributors. Organ Index (entry-point check). https://organindex.de/. Consulted 2026-09-10. bounded_direct_http_get; HTTP 403; page content not verified

78. Organ Index contributors. Organ Index - selected document (descriptive title). https://organindex.de. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

79. MODAVIS Aggregator. aggregator/organindex_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 944cdc3083afff17d453215724a32deb5ea874f18cdb7063c7d88a68f31c6938. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/organindex_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

80. Orgbase / Organ Database. Organ database | Home. http://orgbase.nl/. Consulted 2026-09-10. bounded_direct_http_get

81. Orgbase / Organ Database. Organ database | Information on the database. https://orgbase.nl/info_en.html. Consulted 2026-09-10. bounded_direct_http_get

82. Datenbank-Infosystem (DBIS), Universitätsbibliothek Regensburg. DBIS - Orgeldatabase. https://dbis.ur.de/ALL/resources/105082. Consulted 2026-09-10. bounded_direct_http_get; Resource 105082, Orgeldatabase: descriptive paragraph

83. Orgbase / Organ Database. Organ database | Complete description. https://orgbase.nl/scripts/ogb.exe?database=ob2&%250=2000011&LGE=EN&LIJST=lang. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

84. MODAVIS Aggregator. aggregator/orgbase_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 dc4d35999fb94c47f620157a7d73a916d85c75809aa9243b48f017ea200c4d5c. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/orgbase_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

85. Andreas Schmidt; Stefan Weyreder. Orgel Verzeichnis - Orgelarchiv Schmidt - Deutschland - Schweiz - Österreich. https://orgel-verzeichnis.de/. Consulted 2026-09-10. bounded_direct_http_get

86. Andreas Schmidt; Stefan Weyreder. Über Uns - Orgel Verzeichnis - Orgelarchiv Schmidt. https://orgel-verzeichnis.de/ueber-uns/. Consulted 2026-09-10. bounded_direct_http_get

87. Andreas Schmidt; Stefan Weyreder. Mansfeld - Stadtkirche St. Georg (Talkirche) - Orgel Verzeichnis - Orgelarchiv Schmidt. https://orgel-verzeichnis.de/mansfeld-stadtkirche-st-georg-talkirche/. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

88. MODAVIS Aggregator. aggregator/schmidt_history.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 966a29b46a8f81b967304662311fa06cdfbf896c607572fe9f8e915babccb5fd. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/schmidt_history.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

89. Michael Bernhard (editor); Gesellschaft für Bayerische Musikgeschichte e.V. (original publisher). Orgeldatenbank Bayern. https://orgeldb.de/. Consulted 2026-09-10. bounded_direct_http_get

90. Michael Bernhard (editor); Gesellschaft für Bayerische Musikgeschichte e.V. (original publisher). ODB. https://orgeldb.de/cgi-bin/dispo?satz=917591979. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

91. Hochschule Luzern; Zentral- und Hochschulbibliothek Luzern. Orgeldokumentationszentrum der Hochschule Luzern - Musik. https://orgeldokumentationszentrum.ch/. Consulted 2026-09-10. bounded_direct_http_get

92. Hochschule Luzern; Zentral- und Hochschulbibliothek Luzern. Orgeldokumentationszentrum | Hochschule Luzern. https://www.hslu.ch/de-ch/musik/campus/bibliothek/spezialsammlung-musik/orgeldokumentationszentrum/. Consulted 2026-09-10. bounded_direct_http_get

93. Hochschule Luzern; Zentral- und Hochschulbibliothek Luzern. Orgeldokumentationszentrum HSLU - selected document (descriptive title). https://orgeldokumentationszentrum.ch/backend/v3/details.php?id=17202. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

94. Hochschule für Musik Carl Maria von Weber Dresden. Orgeldatenbank Sachsen - Orgeldatenbank Sachsen. https://www.orgelforum-sachsen.de/orgeldatenbank-sachsen.html. Consulted 2026-09-10. bounded_direct_http_get

95. Hochschule für Musik Carl Maria von Weber Dresden. Ablaß - Orgeldatenbank Sachsen. https://www.orgelforum-sachsen.de/ablass.html?articles=ablass. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

96. MODAVIS Aggregator. aggregator/orgelsachsen_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 9926e666b10eda46e08d1e1a83e611e3d358fadad86410ee20785f7098541690. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/orgelsachsen_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

97. Die Orgelseite. Die Orgelseite - 6.622 Orgeln in 83 Ländern, Fotos, Dispositionen, Musik und CDs. https://die-orgelseite.de/. Consulted 2026-09-10. bounded_direct_http_get

98. Die Orgelseite. Impressum. https://die-orgelseite.de/impressum.htm. Consulted 2026-09-10. bounded_direct_http_get

99. Die Orgelseite. Orgeldatenbank mit Fotos, Dispositionen und Klangbeispielen - Seite 001. https://die-orgelseite.de/orgelliste_001.htm. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

100. MODAVIS Aggregator. aggregator/orgelseite_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 a0a6abf20fcdf54a32e45957cdaa93e373df7016b435d9f8093efe06f8b4e716. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/orgelseite_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

101. de Orgelsite / orgelsite.nl. de Orgelsite | orgelsite.nl. https://www.orgelsite.nl/. Consulted 2026-09-10. bounded_direct_http_get

102. de Orgelsite / orgelsite.nl. Dordrecht, Grote Kerk, Hoofdorgel - de Orgelsite | orgelsite.nl. https://www.orgelsite.nl/dordrecht-grote-kerk-hoofdorgel/. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

103. Peter Fasler; François Comment / Orgelverzeichnis Schweiz. Orgelverzeichnis Schweiz. https://orgelverzeichnis.ch/. Consulted 2026-09-10. bounded_direct_http_get

104. Peter Fasler; François Comment / Orgelverzeichnis Schweiz. Urheber. https://orgelverzeichnis.ch/Orgelverzeichnis/urheber.htm. Consulted 2026-09-10. bounded_direct_http_get

105. Peter Fasler; François Comment / Orgelverzeichnis Schweiz. Aarau alte Kant. Schule, Orgel-Profile Aargau. https://orgelverzeichnis.ch/AGProfile/aarau_gymalt.htm. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

106. Orgues.ch. Orgues.ch - Orgues en Suisse. https://www.orgues.ch/. Consulted 2026-09-10. bounded_direct_http_get

107. Orgues.ch. Orgue non trouvé. https://www.orgues.ch/orgues/orgue-de-eglise-catholique-de-burgdorf. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

108. Orgues en France et dans le monde. Accueil. http://orguesfrance.com/. Consulted 2026-09-10. bounded_direct_http_get

109. Orgues en France et dans le monde. Auditorium de Radio-France - Paris 16è. http://orguesfrance.com/ParisAuditoriumRadioFrance.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

110. Organ Society of Western Australia. The Organ Society of Western Australia. https://www.oswa.org.au/. Consulted 2026-09-10. bounded_direct_http_get

111. Organ Society of Western Australia. Pipe Organs in WA. https://www.oswa.org.au/sup/WAOrgans/Albany/AlbanyFreeReformed.html. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

112. van Ginkel, Gert. PC organ Hauptwerk virtual software organ on computer via midi. Compare sampleset with free mp3. https://www.pcorgan.com/SampleSetsEN.html. Consulted 2026-09-10. bounded_direct_http_get

113. van Ginkel, Gert. PC organ Hauptwerk virtual software organ on computer via midi. Gert van Ginkel The Netherlands. Compare sampleset with free mp3. https://www.pcorgan.com/OverPCorganEN.html. Consulted 2026-09-10. bounded_direct_http_get

114. Piotr Grabowski. GrandOrgue versions - Piotr Grabowski - Virtual Pipe Organ Sample Sets. https://piotrgrabowski.pl/grandorgue-versions/. Consulted 2026-09-10. bounded_direct_http_get

115. Piotr Grabowski. Instruments - Piotr Grabowski - Virtual Pipe Organ Sample Sets. https://piotrgrabowski.pl/instruments/. Consulted 2026-09-10. bounded_direct_http_get

116. Pipedreams. Pipedreams | Pipedreams. https://www.pipedreams.org/. Consulted 2026-09-10. bounded_direct_http_get

117. Pipedreams. 1987 Casavant organ at the Jack Singer Concert Hall of the Calgary Center for the Performing Arts, Calgary, Canada | Pipedreams. https://www.pipedreams.org/profile/gallery/canada/calgary_jack-singer_casavant. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

118. Leif-Erik Hallmann / Pipe Organ Map. Home: Discover the World of Pipe Organs | Pipe Organ Map. https://pipeorganmap.com/. Consulted 2026-09-10. bounded_direct_http_get

119. Leif-Erik Hallmann / Pipe Organ Map. Lille, Cathédrale Notre-Dame-de-la-Treille | Pipe Organ Map. https://pipeorganmap.com/organ/lille-cathédrale-notre-dame-de-la-treille. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

120. RIdIM Germany / Bayerische Staatsbibliothek. RIdIM. https://ridim.musiconn.de/ridim/. Consulted 2026-09-10. bounded_direct_http_get

121. RIdIM Germany / Bayerische Staatsbibliothek. Impressum - https://ridim.musiconn.de. https://ridim.musiconn.de/impressum. Consulted 2026-09-10. bounded_direct_http_get

122. MODAVIS Aggregator. aggregator/ridim_iconography_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 985fa8604c796f3e47266c7254b20e4f0b2734e86020168448ec13e506ee73c9. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/ridim_iconography_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

123. MODAVIS Aggregator. aggregator/ridim_iconography_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 a78db4720d48770b01031616f17c9b9c65da50549f62b67270e645ca96ffaf80. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/ridim_iconography_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

124. Sakralorgelforum. Sakralorgelforum. https://www.sakralorgelforum.net/. Consulted 2026-09-10. bounded_direct_http_get

125. Göteborg International Organ Academy Association; University of Gothenburg / GRIDH. Sonora: Digitising Sweden's Organ Heritage. https://sonora.dh.gu.se/. Consulted 2026-09-10. rendered_browser_page; About the portal; The Archives

126. Göteborg International Organ Academy Association; University of Gothenburg / GRIDH. 404 Not Found. https://organacademy.se/svenskt-digitalt-orgelarkiv. Consulted 2026-09-10. bounded_direct_http_get; HTTP 404; page content not verified

127. Göteborg International Organ Academy Association; University of Gothenburg / GRIDH. SONORA - selected document (descriptive title). https://orgeldatabas.gu.se/webgoart/goart/organ1.php?id=14303&lang=sv. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

128. Sonus Paradisi. Sonus Paradisi Free Stuff / Blog SonusParadisi.cz. https://www.sonusparadisi.cz/en/blog/category/free-stuff/. Consulted 2026-09-10. bounded_direct_http_get

129. GrandOrgue project contributors. GrandOrgue / Samplesets / Sample Sets. https://sourceforge.net/p/ourorgan/samplesets/Sample%20Sets/. Consulted 2026-09-10. bounded_direct_http_get

130. Svoboda, Štěpán. Varhany a varhanáři v České republice. https://www.varhany.net/. Consulted 2026-09-10. bounded_direct_http_get

131. Svoboda, Štěpán. Varhany a varhanáři v České republice. https://www.varhany.net/info.php. Consulted 2026-09-10. bounded_direct_http_get

132. Štěpán Svoboda. Varhany a varhanáři v České republice. https://www.varhany.net/cardheader.php?id=403&lok=268&podrobnosti=1. Consulted 2026-09-11. Selected document inspected or attempted on 11 September 2026; see evidence.json and the individual catalogue account. Access is not a source-wide validation.

133. MODAVIS Aggregator. aggregator/varhany_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 2aa567bfe00ca08330d348c3d4113fba5a0864a8696abd8be4ff796faca7f799. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/varhany_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

134. Frederick Russ / VI-Control. Vi-Control. https://vi-control.net/portal/. Consulted 2026-09-10. bounded_direct_http_get

135. MODAVIS Aggregator. aggregator/vi_control_forum_crawler.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 6bd444406da3d2fbd896b5c3a13a86d6f2e717a2b063d12670fbadb6c2116875. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/vi_control_forum_crawler.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

136. MODAVIS Aggregator. aggregator/vmi_forum_transfer_schema.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 36bd22234798c8458672138a454ad8b316f60f4a5ea79eda01126302360aa1a5. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/vmi_forum_transfer_schema.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

137. Virtual Pipe Organs. Free Sample Sets and Producers - Virtual Pipe Organs. https://virtualpipeorgans.wordpress.com/sample-sets/free-sample-set-producers/. Consulted 2026-09-10. bounded_direct_http_get

138. Virtual Pipe Organs. About - Virtual Pipe Organs. https://virtualpipeorgans.wordpress.com/about/. Consulted 2026-09-10. bounded_direct_http_get

139. Virtual Pipe Organs. ODFs For GrandOrgue - Virtual Pipe Organs. https://virtualpipeorgans.wordpress.com/sample-sets/odfs-for-grandorgue/. Consulted 2026-09-10. bounded_direct_http_get

140. Long, Phillip; Novack, Zachary; McAuley, Julian; Berg-Kirkpatrick, Taylor. PDMX: A Large-Scale Public Domain MusicXML Dataset for Symbolic Music Processing | Zenodo. 2025-6-1. Version v9. https://zenodo.org/records/15571083. Consulted 2026-09-10. bounded_direct_http_get

141. Asai, H.; Gao, Y.; Shimosaka, S.; Tojo, S.; Uehara, Y.. GitHub - yui-u/little-organ-analyses: The Little Organ Book Dataset · GitHub. https://github.com/yui-u/little-organ-analyses. Consulted 2026-09-10. bounded_direct_http_get

142. IMSLP / Petrucci Music Library; Werner Icking Music Archive contributors. IMSLP: Free Sheet Music PDF Download. https://imslp.org/. Consulted 2026-09-10. bounded_direct_http_get

143. The Mutopia Project and individual typesetters. The Mutopia Project. https://www.mutopiaproject.org/. Consulted 2026-09-10. bounded_direct_http_get

144. The Mutopia Project and individual typesetters. Mutopia: License details. https://www.mutopiaproject.org/legal.html. Consulted 2026-09-10. bounded_direct_http_get

145. MuseScore.com and individual score contributors. MuseScore.com and individual score contributors (entry-point check). https://musescore.com/. Consulted 2026-09-10. bounded_direct_http_get; HTTP 403; page content not verified

146. Evangelisch-Lutherische Landeskirche Sachsens. EVLKS - interessiert: Orgeln. https://www.evlks.de/feiern/kirchenmusik/orgeln/#c3681. Consulted 2026-09-10. bounded_direct_http_get

147. MODAVIS project. MODAVIS Aggregator: sources. Version 2bb6e85282e3ab11af1bcc7126290f67af176f53. Packaged extract in evidence.json; source path documentation/source_investigation_audit.yaml; SHA-256 9713e7fe5625204fd806fb7fe9c670e3af84827e1b3a2d091b2110b2744053e7

148. Stauff, Edward L.. Encyclopedia of Organ Stops. http://www.organstops.org/. Consulted 2026-09-10. bounded_direct_http_get

149. Welt-Orgel-Archiv (title recorded in the investigation audit). Welt-Orgel-Archiv (title recorded in the investigation audit) (entry-point check). http://orgelarchiv.orgel-information.de/. Consulted 2026-09-10. bounded_direct_http_get; HTTP 404; page content not verified

150. MODAVIS Aggregator. aggregator/indo_event_participants.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 743a8df6674bf64c784af4583952731f9ff7c241f869335997260a17cedfce3e. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/indo_event_participants.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

151. MODAVIS Aggregator. aggregator/swiss_specification_accounts.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 2dac0c6d3e0d1ff0ce63a7666fa82af87aa91c5d9f56d32e240e46b30c5541af. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/swiss_specification_accounts.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.

152. MODAVIS Aggregator. aggregator/swiss_configuration_date.py. Revision 2bb6e85282e3ab11af1bcc7126290f67af176f53; inspected locally 11 September 2026. SHA-256 39389db65a663526d27799ce5b5467d57f1a93e1f19ed82c4113b004aabce680. Repository locator: https://github.com/modavis-project/modavis-aggregator/blob/2bb6e85282e3ab11af1bcc7126290f67af176f53/aggregator/swiss_configuration_date.py. Public access to the tested pinned repository link was not established; see the deposited implementation evidence.