

Package: starc (via r-universe)

October 7, 2026

Title A Provider-Agnostic STAC Reference Cache

Version 0.2.0

Description Harvests SpatioTemporal Asset Catalog (STAC) discovery results into an append-only Parquet reference store: queries, acquisitions, products, and assets as separate levels, with all asset references recorded verbatim and no fixed band vocabulary. Markers and coverage are derived views, never stored state. Stores references, never pixels.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

Depends R (>= 4.2.0)

Imports arrow, digest, dplyr, jsonlite, sds

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

Config/roxygen2/version 8.1.0

Config/pak/sysreqs cmake libssl-dev

Repository <https://hypertidy.r-universe.dev>

Date/Publication 2026-10-06 12:34:12 UTC

RemoteUrl <https://github.com/hypertidy/starc>

RemoteRef HEAD

RemoteSha 810d1d0a9d6a22ef6a670f7147e84bb3b7b300da

Contents

acquisition_id	2
fetch_all_pages	2
harvest	3
harvest_mgrs	3
mappers	4

marker	4
marker_region	5
mgrs_month_url	5
pad_extent	6
solarday_at	6
tiles_to_mgrs	6
Index	8

acquisition_id	<i>Canonical acquisition id</i>
----------------	---------------------------------

Description

Sentinel-2: platform + MGRS tile + datetime-to-seconds. Survives the Element84 / Planetary Computer / DEA seam; reprocessings are the same acquisition, different product rows. SAR mappers will supply burst/frame in tile.

Usage

acquisition_id(ident)

fetch_all_pages	<i>Follow rel="next" links, collecting all items</i>
-----------------	--

Description

Follow rel="next" links, collecting all items

Usage

fetch_all_pages(url, max_pages = 100)

Arguments

url	First-page search URL (from sds::stacit())
max_pages	Safety valve

Value

list(items = list of item lists, n_pages = int)

harvest	<i>Harvest a region bbox into the reference store</i>
---------	---

Description

Harvest a region bbox into the reference store

Usage

```
harvest(  
  region,  
  provider,  
  collection,  
  t0,  
  t1,  
  store = "~/starc-store",  
  limit = 300,  
  pad_km = 5  
)
```

Arguments

region	One row of a regions table (region_id, lonmin, lonmax, latmin, latmax)
provider	sds::stacit provider token
collection	Collection id (must have a registered mapper)
t0, t1	Query window (RFC3339 or dates sds accepts)
store	Root directory of the parquet store
limit	Page size
pad_km	Discovery buffer applied to the extent (see pad_extent)

Value

the query log row, invisibly

harvest_mgrs	<i>Harvest one (MGRS tile, month) into the reference store</i>
--------------	--

Description

The deterministic sibling of harvest(): same store, same tables, same query-log-last discipline. region_id in the query log is "mgrs:", keeping tile-month harvests distinguishable from region-bbox harvests while sharing every derived view.

Usage

```
harvest_mgrs(  
  mgrs,  
  month,  
  collection = "sentinel-2-c1-l2a",  
  endpoint = "https://earth-search.aws.element84.com/v1/search",  
  store = "~/starc-store",  
  limit = 300  
)
```

Arguments

mgrs	MGRS code, e.g. "43DDE" (no "MGRS-" prefix)
month	"YYYY-MM"
collection	Collection id
endpoint	STAC search endpoint (raw URL; this mode constructs the query itself rather than going through sds::stacit)
store	Root directory of the parquet store
limit	Page size

Value

the query log row, invisibly

mappers	<i>Mapper registry</i>
---------	------------------------

Description

A mapper declares product_family and two functions: identity() returns the fields that name the physical acquisition across providers; promote() returns properties promoted to product columns. Everything else in the item is preserved via the raw stash and the properties JSON column.

Usage

```
mappers
```

marker	<i>Marker: max solarday per collection (deduped, derived, never stored)</i>
--------	---

Description

Marker: max solarday per collection (deduped, derived, never stored)

Usage

```
marker(store = "~/starc-store")
```

marker_region	<i>Region-scoped marker: what estinel's incremental cadence asks for</i>
---------------	--

Description

Joins products -> queries via query_id to recover region_id. Product rows from harvests predating the query_id column drop out of the inner join (re-harvest to backfill).

Usage

```
marker_region(store = "~/starc-store")
```

mgrs_month_url	<i>Build a grid:code + month search URL</i>
----------------	---

Description

Build a grid:code + month search URL

Usage

```
mgrs_month_url(  
  mgrs,  
  month,  
  collection = "sentinel-2-c1-l2a",  
  endpoint = "https://earth-search.aws.element84.com/v1/search",  
  limit = 300  
)
```

Arguments

mgrs	MGRS code, e.g. "43DDE" (no "MGRS-" prefix)
month	"YYYY-MM"
collection	Collection id
endpoint	STAC search endpoint (raw URL; this mode constructs the query itself rather than going through sds::stacit)
limit	Page size

pad_extent	<i>Pad a lonlat extent outward by a minimum ground distance</i>
------------	---

Description

Region extents are member-centroid bboxes, so singleton and co-located regions degenerate to points; discovery queries need area. Pads by min_km on every side, longitude scaled by cos(lat) so the pad is honest at 77S. This is a discovery buffer, not the site window: scene footprints are ~100 km, so modest padding suffices to catch every scene touching the sites.

Usage

```
pad_extent(extent, min_km = 5)
```

solarday_at	<i>Solar day at the item centroid</i>
-------------	---------------------------------------

Description

Solar day at the item centroid

Usage

```
solarday_at(datetime, centroid_lon)
```

tiles_to_mgrs	<i>Map aatgrid tiles to intersecting Sentinel-2 MGRS scenes</i>
---------------	---

Description

Same-zone pairs intersect by interval arithmetic (both rectangles in one CRS). Cross-zone pairs reproject the buffered scene extent into the tile's zone via densified project_extent() – the corner-only trap applies to MGRS squares exactly as it did to Heard's bbox.

Usage

```
tiles_to_mgrs(tiles, valid, lat_bands)
```

Arguments

- tiles data.frame from tiles_for_extent2() (tile_id, zone_id, res, col, row)
- valid Character vector of valid S2 MGRS codes (see starc's inst/extdata/valid_mgrs_sentinel2.parquet)
- lat_bands Latitude band letters to consider (derive from the tiles' latitude range; e.g. c("C", "D", "E") around 53S-77S)

Value

long data.frame: tile_id, mgrs, same_zone, offset60 (scene-origin easting mod 60 in the tile's CRS frame – whether 60 m assets land lattice-aligned or sub-pixel shifted)

Index

acquisition_id, [2](#)
fetch_all_pages, [2](#)
harvest, [3](#)
harvest_mgrs, [3](#)
mappers, [4](#)
marker, [4](#)
marker_region, [5](#)
mgrs_month_url, [5](#)
pad_extent, [6](#)
solarday_at, [6](#)
tiles_to_mgrs, [6](#)