Package 'scholidonline'

June 14, 2026

Type Package

Title Resolution, Conversion, Linking and Metadata for Scholarly Identifiers

Version 0.2.0

Language en-US

Description Enables querying of scholarly identifier services to verify identifier existence, convert identifiers across systems, retrieve bibliographic metadata, and discover linked identifiers. Supports identifier types including DOI, PMID, PMCID, arXiv, ORCID, OpenAlex, ROR, UniProt, and selected NCBI accessions (GEO, BioProject, RefSeq, SRA, and genome assembly).

License MIT + file LICENSE

URL https://thomas-rauter.github.io/scholidonline/

BugReports https://github.com/Thomas-Rauter/scholidonline/issues

Depends R (>= 4.0.0)

Imports scholid (>= 0.2.0), httr2, rlang

Suggests testthat (>= 3.0.0), knitr (>= 1.30), rmarkdown

Encoding UTF-8

RoxygenNote 7.3.3

Config/testthat/edition 3

VignetteBuilder knitr

NeedsCompilation no

Author Thomas Rauter [aut, cre, fnd] (ORCID: <https://orcid.org/0009-0004-5578-3628>)

Maintainer Thomas Rauter <rauterthomas0@gmail.com>

Repository CRAN

Date/Publication 2026-06-14 17:50:02 UTC

Contents

`id_convert` *Convert scholarly identifiers across systems*

Description

Convert scholarly identifiers across registries, for example from PMID to DOI.

Usage

```
id_convert(
  x,
  to = scholidonline_types(),
  from = NULL,
  provider = c("auto", .scholidonline_providers()),
  ...,
  quiet = FALSE
)
```

Arguments

`x`	A character vector of scholarly identifiers.
`to`	A single target identifier type string, such as `"doi"` or `"pmid"`. See `scholidonline_types()` for all supported values.
`from`	A single source identifier type string, or `NULL` to infer the source type for each element of `x`.
`provider`	A single provider string specifying which online service to use for the conversion. Use `"auto"` to use the default provider for the requested conversion. In most cases, `"auto"` is appropriate.
`...`	Reserved for future provider-specific arguments.
`quiet`	A single logical value; if `TRUE`, suppress provider warnings and messages where possible.

Details

Only some source/target type pairs are supported. Use `scholidonline_capabilities()` with `operation = "convert"` (or filter the returned table) to see which conversions are available and which providers implement them.

Value

A character vector of converted identifiers. Elements that cannot be identified, normalized, or converted return `NA_character_`.

Examples

```
id_convert("12345678", to = "doi", from = "pmid")
id_convert("10.1038/nature12373", to = "pmid", from = "doi")
```

`id_exists` *Check whether scholarly identifiers exist*

Description

Check whether scholarly identifiers are found in their respective registries.

Usage

```
id_exists(
  x,
  type = c("auto", scholidonline_types()),
  provider = c("auto", .scholidonline_providers()),
  ...,
  quiet = FALSE
)
```

Arguments

`x`	A character vector of identifiers.
`type`	A single identifier type string, or `"auto"` to infer the type for each element of `x`. See `scholidonline_types()` for supported values.
`provider`	A single provider string specifying which online service to use for the lookup. Use `"auto"` to use the default provider for the resolved identifier type. In most cases, `"auto"` is appropriate.
`...`	Reserved for future provider-specific arguments.
`quiet`	A single logical value; if `TRUE`, suppress provider warnings and messages where possible.

Details

Existence checking is not available for every identifier type supported by `scholid`. Use `scholidonline_capabilities()` to see which types support the `exists` operation and which providers implement it.

`type` must be a single value or `"auto"`. For mixed identifier columns, omit `type` or use `type = "auto"` so each element is classified separately.

Value

A logical vector. TRUE indicates that the identifier was found, FALSE indicates that it was not found, and NA indicates that the input could not be identified, normalized, or checked reliably.

Examples

```
id_exists("10.1038/nature12373", type = "doi")
id_exists(c("31452104", "PMC6784763"))
```

---

id_links *Return identifiers linked to the same scholarly record*

Description

Return identifiers that external registries link to the same scholarly record or to a closely corresponding version of it.

Usage

```
id_links(
  x,
  type = c("auto", scholidonline_types()),
  provider = c("auto", .scholidonline_providers()),
  ...,
  quiet = FALSE
)
```

Arguments

x	A character vector of identifiers.
type	A single identifier type string, or "auto" to infer the type for each element of x. See scholidonline_types() for supported values.
provider	A single provider string specifying which online service to use. Use "auto" to use the default provider for the resolved identifier type. In most cases, "auto" is appropriate.
...	Reserved for future provider-specific arguments.
quiet	A single logical value; if TRUE, suppress provider warnings and messages where possible.

Details

`id_links()` is vectorized over `x` and returns a long data.frame with one row per discovered identifier link.

Typical links include DOI <-> PMID, DOI <-> PMCID, PMID <-> PMCID, arXiv ID <-> DOI, ORCID -> DOI for works recorded in ORCID, and OpenAlex work -> DOI, PMID, or PMCID where present in the OpenAlex record.

Link discovery is not available for every supported identifier type; use `scholidonline_capabilities()` to check whether `links` is supported.

Only identifier links explicitly exposed by the queried provider are returned. `id_links()` does not retrieve general metadata or broader related records unless the provider represents them as direct identifier links.

Trivial self-links are excluded from the result.

`type` must be a single value or `"auto"`. For mixed identifier columns, omit `type` or use `type = "auto"` so each element is classified separately.

Value

A data.frame with columns `query`, `query_type`, `linked_type`, `linked_id`, and `provider`. If no links are found, a zero-row data.frame with these columns is returned.

Examples

```
out <- id_links("31452104", provider = "epmc")
knitr::kable(out)
```

`id_metadata` *Retrieve scholarly metadata*

Description

Retrieve structured metadata for scholarly identifiers from external registries.

Usage

```
id_metadata(
  x,
  type = c("auto", scholidonline_types()),
  provider = c("auto", .scholidonline_providers()),
  fields = NULL,
  ...,
  quiet = FALSE
)
```

Arguments

x	A character vector of identifiers.
type	A single identifier type string, or "auto" to infer the type for each element of x. See scholidonline_types() for supported values.
provider	A single provider string specifying which online service to use. Use "auto" to use the default provider for the resolved identifier type. In most cases, "auto" is appropriate.
fields	An optional character vector naming the columns to return. If NULL, all default columns are returned. Unknown field names are ignored.
...	Reserved for future provider-specific arguments.
quiet	A single logical value; if TRUE, suppress provider warnings and messages where possible.

Details

id_metadata() is vectorized over x and returns a data.frame with one row per input identifier.

For providers that support batch lookup, such as arXiv, multiple identifiers may be resolved using a single provider request. This does not change the public return shape: the output still contains one row per input identifier.

The function returns a harmonized cross-provider data.frame with columns title, year, container, doi, pmid, pmcid, and url. For bibliographic identifiers, container is typically a journal or source title and linked DOI/PMID/PMCID fields may be populated. For other types, the same columns are reused with type-appropriate meaning (for example, protein name and organism for UniProt, organization name and country for ROR, or accession title and organism for NCBI accessions). Bibliographic link columns are NA when not applicable.

For NCBI accession types such as BioProject, title is the registry's short project or record title from Entrez ESummary, not the full description shown on the NCBI website. Use url for the complete record.

type must be a single value or "auto". For mixed identifier columns, omit type or use type = "auto" so each element is classified separately.

Value

A data.frame with one row per input identifier. By default, the returned columns are input, type, provider, title, year, container, doi, pmid, pmcid, and url. Inputs that cannot be identified, normalized, or resolved are returned as rows with missing metadata fields.

Examples

```
out <- id_metadata("10.1038/nature12373", type = "doi")
knitr::kable(out)
out <- id_metadata(c("31452104", "PMC6821181"))
knitr::kable(out)
out <- id_metadata(
  "10.1038/nature12373",
   fields = c("title", "year", "doi")
```

```
  )
knitr::kable(out)
```

`scholidonline_capabilities` *Supported scholidonline capabilities*

Description

Return a summary of the capabilities supported by the scholidonline package.

The returned table describes, for each supported identifier type:

- which single-identifier operations are available (`exists`, `links`, `meta`),
- which identifier conversions are available,
- which providers support each capability, and
- which provider is used by default when `provider = "auto"`.

This function is useful for discovering what scholidonline can do for a given identifier type or conversion pair.

Usage

```
scholidonline_capabilities()
```

Value

A data.frame with one row per supported capability and the following columns:

- `type`: source identifier type
- `operation`: operation name (`exists`, `links`, `meta`, or `convert`)
- `target`: target identifier type for conversion operations, otherwise `NA`
- `providers`: comma-separated names of providers supporting the capability
- `default_provider`: default provider used when `provider = "auto"`

Examples

```
caps <- scholidonline_capabilities()

subset(caps, type == "pmid" & operation == "convert")

subset(caps, type == "doi" & target == "pmcid")
```

`scholidonline_types` *Supported scholidonline identifier types*

Description

Return the set of identifier types supported by the scholidonline package.

This is the set of identifier types for which scholidonline provides registry-backed functionality. Available operations vary by type; use `scholidonline_capabilities()` to see which of existence checks, metadata retrieval, link discovery, and identifier conversion are supported for each type.

Usage

```
scholidonline_types()
```

Value

A character vector of supported identifier type strings.

Examples

```
scholidonline_types()
"doi" %in% scholidonline_types()
```

Index