Package 'xxdi'

August 22, 2026

Type Package

Title An R Package for Evaluating Expertise Indices for Research Strength Assessment

Version 1.26.8

Maintainer Nilabhra R Das <nr.das@yahoo.com>

Description Institutional performance assessment remains a key challenge to a multitude of stakeholders. Existing indicators such as h-type indicators, g-type indicators, and many others do not reflect expertise of institutions that defines their research portfolio. The package offers functionality to compute and visualise two novel indices: the x-index and the xd-index. The x-index evaluates an institution's scholarly expertise within a specific discipline or field, while the xd-index provides a broader assessment of overall scholarly expertise considering an institution's publication pattern and strengths across coarse thematic areas. These indices offer a nuanced understanding of institutional research capabilities, aiding stakeholders in research management and resource allocation decisions. For further details, see Lathabai, H.H., Nandy, A., and Singh, V.K. (2021) <doi:10.1007/s11192-021-04188-3>, and Nandy, A., Lathabai, H.H., and Singh, V.K. (2023) <doi:10.5281/zenodo.8305585>. This package provides wrapper functions for computing these exertise indices and their variants, for use with the standard output format of the Web of Science database.

License GPL-3

Depends R (>= 4.4.2.0)

Imports agop, checkmate, dplyr, ggplot2, stats, tidyr

Encoding UTF-8

RoxygenNote 7.3.3

LazyData true

NeedsCompilation no

Author Nilabhra R Das [cre, aut] (ORCID: <https://orcid.org/0000-0001-8187-0080>), Abhirup Nandy [aut] (ORCID: <https://orcid.org/0000-0001-8618-0847>)

Repository CRAN

Date/Publication 2026-08-22 15:20:02 UTC

Contents

g_index 2
h_index 3
ivw_xd_index 4
WoSdata 5
xc_index 6
xd_index 7
xo_index 9
x_index 10

Index **12**

`g_index` *g_index - Egghe's g-index*

Description

Calculate g-index for an institution using bibliometric data from an edge list, with an optional visualisation of ranked citation scores.

Usage

```
g_index(df, id = NULL, cit, plot = FALSE)
```

Arguments

`df`	Data frame object containing bibliometric data. This data frame must have at least two columns: one for keywords and one for citation counts. An optional column for unique identifiers can be included. Each row in the data frame should represent a document or publication.
`id`	Character string specifying the name of the column in "df" that contains unique identifiers for each document. Each cell in this column must contain a single ID (unless missing) and not multiple IDs. Must be included when 'plot' parameter is set to "TRUE". Default set to NULL.
`cit`	Character string specifying the name of the column in "df" that contains the number of citations each document has received. Citations must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the citation count for the corresponding document.
`plot`	Logical value indicating whether to generate and display a plot of the g-index calculation. Set to "TRUE" or "T" to generate the plot, and "FALSE" (default) or "F" to skip plot generation.

Value

g-index magnitude and optional plot.

Examples

```
# Load example data
data(WoSdata)

# Calculate g-index with plot
g_index(df = WoSdata,
        id = "UT.Unique.WOS.ID",
        cit = "Times.Cited.WoS.Core",
        plot = TRUE)
```

---

h_index *h_index - Hirsch's h-index*

Description

Calculate h-index for an institution using bibliometric data from an edge list, with an optional visualisation of ranked citation scores.

Usage

```
h_index(df, id = NULL, cit, plot = FALSE)
```

Arguments

df	Data frame object containing bibliometric data. This data frame must have at least two columns: one for keywords and one for citation counts. An optional column for unique identifiers may be included. Each row in the data frame should represent a document or publication.
id	Character string specifying the name of the column in "df" that contains unique identifiers for each document. Each cell in this column must contain a single ID (unless missing) and not multiple IDs. Must be included when 'plot' parameter is set to "TRUE". Default set to NULL.
cit	Character string specifying the name of the column in "df" that contains the number of citations each document has received. Citations must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the citation count for the corresponding document.
plot	Logical value indicating whether to generate and display a plot of the h-index calculation. Set to "TRUE" or "T" to generate the plot, and "FALSE" (default) or "F" to skip plot generation.

Value

h-index magnitude and optional plot.

Examples

```
# Load example data
data(WoSdata)

# Calculate h-index with produce plot
h_index(df = WoSdata,
        id = "UT.Unique.WOS.ID",
        cit = "Times.Cited.WoS.Core",
        plot = TRUE)
```

`ivw_xd_index` *ivw_xd_index - Inverse Variance Weighted (IVW) Expertise Diversity Index*

Description

Calculate IVW adjusted xd-index for an institution using bibliometric data from an edge list, with an optional visualisation of ranked citation scores. The function is suitable for including inside loops when plotting parameter is set to "FALSE" or "F".

Usage

```
ivw_xd_index(df, cat, id, cit, vfc = NULL, type = "h", dlm = ";", plot = FALSE)
```

Arguments

`df`	Data frame object containing bibliometric data. This data frame must have at least three columns: one for categories, one for unique IDs, and one for citation counts. Each row in the data frame should represent a document or publication.
`cat`	Character string specifying the name of the column in "df" that contains categories. Each cell in this column may contain no categories (missing), a single category or multiple categories separated by a specified delimiter.
`id`	Character string specifying the name of the column in "df" that contains unique identifiers for each document. Each cell in this column must contain a single ID (unless missing) and not multiple IDs.
`cit`	Character string specifying the name of the column in "df" that contains the number of citations each document has received. Citations must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the citation count for the corresponding document.
`vfc`	Data frame with columns 'cat' and 'var'. Optionally required for using population variances. If not provided estimated sample variances will be used.
`type`	"h" (default) for Hirsch's h-type index or "g" for Egghe's g-type index.
`dlm`	Character string specifying the delimiter used in the "cat" column to separate multiple categories within a single cell. The delimiter should be consistent across the entire "cat" column. Common delimiters include ";" (default), "/", ":", and ",".

`plot` Logical value indicating whether to generate and display a plot of the xd-index calculation. Set to "TRUE" or "T" to generate the plot, and "FALSE" (default) or "F" to skip plot generation.

Value

IVW xd-index magnitude, core categories, and optional plot.

Examples

```
# Load example data
data(WoSdata)

# Calculate ivw xd-index with plot
ivw_xd_index(df = WoSdata,
        id = "UT.Unique.WOS.ID",
        cat = "WoS.Categories",
        cit = "Times.Cited.WoS.Core",
        plot = TRUE)
```

`WoSdata` *WoSdata*

Description

The list of publications and associated metadata for a scholarly institution in India queried from the Web of Science (WoS) database. All publications, totalling 2,355 distinct publications, within a time frame of 10 years, spanning from 2011 to 2020 were extracted. For these publications, additional information/metadata, namely the 'UT (Unique WoS ID)', 'Keywords Plus', 'WoS Categories', and 'Times Cited, WoS Core' fields, were also extracted.

Usage

```
WoSdata
```

Format

A data frame with 2,355 rows and 4 columns. Each row represents a unique publication:

UT.Unique.WOS.ID Unique publication identifier.

Keywords.Plus Indexed keywords separated by ';'s.

WoS.Categories Indexed categories separated by ';'s.

Times.Cited.WoS.Core Total citations as recorded in the WoS database.

Source

* [WoS](https://clarivate.com/academia-government/scientific-and-academic-research/research-discovery-and-referencing/web-of-science/)

xc_index *xc_index - Category-adjusted Expertise Index*

Description

Calculate the xc-index for an institution using bibliometric data from an edge list, with an optional plot visualisation. The function is suitable for including inside loops when plotting parameter is set to "FALSE" or "F".

Usage

```
xc_index(df, kw, cat, id, cit, type = "h", dlm = c(";", ";"), plot = FALSE)
```

Arguments

df	Data frame object containing bibliometric data. This data frame must have at least three columns: one for keywords, one for unique IDs, and one for citation counts. Each row in the data frame should represent a document or publication.
kw	Character string specifying the name of the column in "df" that contains keywords. Each cell in this column may contain no keywords (missing), a single keyword or multiple keywords separated by a specified delimiter.
cat	Character string specifying the name of the column in "df" that contains categories. Each cell in this column may contain no categories (missing), a single category or multiple categories separated by a specified delimiter.
id	Character string specifying the name of the column in "df" that contains unique identifiers for each document. Each cell in this column must contain a single ID (unless missing) and not multiple IDs.
cit	Character string specifying the name of the column in "df" that contains the number of citations each document has received. Citations must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the citation count for the corresponding document.
type	"h" (default) for Hirsch's h-type index or "g" for Egghe's g-type index.
dlm	Character vector specifying the delimiter used in the "kw" and "cat" columns to separate multiple keywords and categories within a single cell. The delimiter should be consistent across the entirety of the two columns. The first element in the vector is used as the delimiter for keywords, while the second element is used for categories. Common delimiters include ";", "/", ":", and ",", and example usage should be similar to 'c(";", ";")' (default), 'c(";", ",")', or 'c(",", ":")'.
plot	Logical value indicating whether to generate and display a plot of the xc-index calculation. Set to "TRUE" or "T" to generate the plot, and "FALSE" (default) or "F" to skip plot generation.

Value

xc-index magnitude, core category specific keywords and optional plot.

Examples

```
# Load example data
data(WoSdata)

# Calculate xc-index with plot
xc_index(df = WoSdata,
         id = "UT.Unique.WOS.ID",
         kw = "Keywords.Plus",
         cat = "WoS.Categories",
         cit = "Times.Cited.WoS.Core",
         plot = TRUE)
```

---

xd_index *xd_index - Expertise Diversity Index*

Description

Calculate the xd-index (and its variants, field-normalized and fractional) for an institution using bibliometric data from an edge list, with an optional visualisation of ranked citation scores.

Usage

```
xd_index(
  df,
  cat,
  id,
  cit,
  inst = NULL,
  mfc = NULL,
  type = "h",
  dlm = ";",
  variant = "U",
  plot = FALSE
)
```

Arguments

df	Data frame object containing bibliometric data. Must have at least three columns: one for categories, one for unique IDs, and one for citation counts.
cat	Character string specifying the name of the column in "df" that contains categories. Categories can be multiple separated by a delimiter.
id	Character string specifying the name of the column in "df" that contains unique identifiers for each document. Each cell in this column must contain a single ID (unless missing) and not multiple IDs.

`cit` Character string specifying the name of the column in "df" that contains the number of citations each document has received. Citations must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the citation count for the corresponding document.

`inst` Character string specifying the name of the column in "df" that contains the number of institutions that contributed to each document. The number of institutions must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the institution count for the corresponding document.

`mfc` Data frame with columns 'cat' and 'mean'. Optionally required to utilise population means when variant set to "f".

`type` "h" (default) for Hirsch's h-type index or "g" for Egghe's g-type index.

`dlm` Character string specifying the delimiter used in the "cat" column to separate multiple categories within a single cell. The delimiter should be consistent across the entire "cat" column. Common delimiters include ";" (default), "/", ":", and ",".

`variant` One of "U" (default) for the standard unconditional variant, "f" for the fractional variant, and "FN" for the field-normalised variant.

- `"U"` — Computes the unconditional xd-index.
- `"f"` — Computes the fractional xd-index. If set to 'f', input data frame 'df' must include an 'inst_count' column which gives the number of institutions per publication.
- `"FN"` — Computes the field-normalised xd-index. If set to 'FN', input may optionally include an 'mfc' data frame which gives the population level mean citations for different fields. If not provided, sample mean field citations will be used.

`plot` Logical value indicating whether to generate and display a plot of the xd-index calculation. Set to "TRUE" or "T" to generate the plot, and "FALSE" (default) or "F" to skip plot generation.

Value

xd-index magnitude, core categories, and optional plot.

Examples

```
# Load example data
data(WoSdata)

# Calculate xd-index with plot
xd_index(df = WoSdata,
         id = "UT.Unique.WOS.ID",
         cat = "WoS.Categories",
         cit = "Times.Cited.WoS.Core",
         plot = TRUE)

# Calculate field-normalised xd-index with plot
```

```
xd_index(df = WoSdata,
         id = "UT.Unique.WOS.ID",
         cat = "WoS.Categories",
         cit = "Times.Cited.WoS.Core",
         variant = "FN",
         plot = TRUE)
```

---

xo_index *xo_index - Overall Expertise Index*

Description

Calculate the xo-index for an institution using bibliometric data from an edge list, with an optional plot visualisation. The function is suitable for including inside loops when plotting parameter is set to "FALSE" or "F".

Usage

```
xo_index(df, kw, cat, id, cit, type = "h", dlm = c(";", ";"), plot = FALSE)
```

Arguments

df	Data frame object containing bibliometric data. This data frame must have at least three columns: one for keywords, one for unique IDs, and one for citation counts. Each row in the data frame should represent a document or publication.
kw	Character string specifying the name of the column in "df" that contains keywords. Each cell in this column may contain no keywords (missing), a single keyword or multiple keywords separated by a specified delimiter.
cat	Character string specifying the name of the column in "df" that contains categories. Each cell in this column may contain no categories (missing), a single category or multiple categories separated by a specified delimiter.
id	Character string specifying the name of the column in "df" that contains unique identifiers for each document. Each cell in this column must contain a single ID (unless missing) and not multiple IDs.
cit	Character string specifying the name of the column in "df" that contains the number of citations each document has received. Citations must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the citation count for the corresponding document.
type	"h" (default) for Hirsch's h-type index or "g" for Egghe's g-type index.
dlm	Character vector specifying the delimiter used in the "kw" and "cat" columns to separate multiple keywords and categories within a single cell. The delimiter should be consistent across the entirety of the two columns. The first element in the vector is used as the delimiter for keywords, while the second element is used for categories. Common delimiters include ";", "/", ":", and ",", and example usage should be similar to 'c(";", ";")' (default), 'c(";", ",")', or 'c(",", ":")'.

plot — Logical value indicating whether to generate and display a plot of the xo-index calculation. Set to "TRUE" or "T" to generate the plot, and "FALSE" (default) or "F" to skip plot generation.

Value

xo-index magnitude, core categories and optional plot.

Examples

```
# Load example data
data(WoSdata)

# Calculate xo-index with plot
xo_index(df = WoSdata,
         id = "UT.Unique.WOS.ID",
         kw = "Keywords.Plus",
         cat = "WoS.Categories",
         cit = "Times.Cited.WoS.Core",
         plot = TRUE)
```

---

x_index *x_index - Expertise Index*

Description

Calculate x-index for an institution using bibliometric data from an edge list, with an optional visualisation of ranked citation scores. The function is suitable for including inside loops when plotting parameter is set to "FALSE" or "F".

Usage

```
x_index(df, kw, id, cit, type = "h", dlm = ";", plot = FALSE)
```

Arguments

df — Data frame object containing bibliometric data. This data frame must have at least three columns: one for keywords, one for unique IDs, and one for citation counts. Each row in the data frame should represent a document or publication.

kw — Character string specifying the name of the column in "df" that contains keywords. Each cell in this column may contain no keywords (missing), a single keyword or multiple keywords separated by a specified delimiter.

id — Character string specifying the name of the column in "df" that contains unique identifiers for each document. Each cell in this column must contain a single ID (unless missing) and not multiple IDs.

cit	Character string specifying the name of the column in "df" that contains the number of citations each document has received. Citations must be represented as integers. Each cell in this column should contain a single integer value (unless missing) representing the citation count for the corresponding document.
type	"h" (default) for Hirsch's h-type index or "g" for Egghe's g-type index.
dlm	Character string specifying the delimiter used in the "kw" column to separate multiple keywords within a single cell. The delimiter should be consistent across the entire "kw" column. Common delimiters include ";" (default), "/", ":", and ",".
plot	Logical value indicating whether to generate and display a plot of the x-index calculation. Set to "TRUE" or "T" to generate the plot, and "FALSE" (default) or "F" to skip plot generation.

Value

x-index magnitude, core keywords, and optional plot.

Examples

```
# Load example data
data(WoSdata)

# Calculate x-index with plot
x_index(df = WoSdata,
        id = "UT.Unique.WOS.ID",
        kw = "Keywords.Plus",
        cit = "Times.Cited.WoS.Core",
        plot = TRUE)
```

Index

∗ **datasets**
    WoSdata, 5

g_index, 2

h_index, 3

ivw_xd_index, 4

WoSdata, 5

x_index, 10
xc_index, 6
xd_index, 7
xo_index, 9