

Package ‘xxdi’

June 16, 2025

Type Package

Title Calculate Expertise Indices

Version 1.2.4

Maintainer Nilabhra Rohan 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. Lathabai, H.H., Nandy, A., and Singh, V.K. (2021) <[doi:10.1007/s11192-021-04188-3](https://doi.org/10.1007/s11192-021-04188-3)>. Nandy, A., Lathabai, H.H., and Singh, V.K. (2023) <[doi:10.5281/zenodo.8305585](https://doi.org/10.5281/zenodo.8305585)>. This package provides the h, g, x, and xd indices for use with standard format of Web of Science (WoS) scrapped datasets.

License GPL-3

Depends R (>= 4.3.3.0)

Imports agop (>= 0.2.4), dplyr (>= 1.1.4), ggplot2 (>= 3.5.0), Matrix (>= 1.6.1.1), stats (>= 4.3.3), tidyr (>= 1.3.1)

Encoding UTF-8

RoxygenNote 7.3.2

NeedsCompilation no

Author Nilabhra Rohan 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 2025-06-16 10:20:02 UTC

Contents

g_index	2
h_index	3
xd_index	4
x_index	6
Index	8

g_index	<i>g_index</i>
---------	----------------

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. Only required 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" or "F" to skip plot generation. The default is "FALSE".

Value

g-index value and plot.

Examples

```
# Create an example data frame
dat1 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a; b; c", "b; d", "c", "d", "e; g", "f", "g"),
  id = c("abc123", "bcd234", "def345", "efg456", "fgh567", "ghi678", "hij789"),
  categories = c("a; d; e", "b", "c", "d; g", "e", "f", "g"))

# Calculate g-index
g_index(df = dat1, cit = "citations")

# Create another example data frame
dat2 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a/ b/ c", "b/ d", "c", "d", "e/ g", "f", "g"),
  id = c("123", "234", "345", "456", "567", "678", "789"),
  categories = c("a/ d/ e", "b", "c", "d/ g", "e", "f", "g"))

# Calculate g-index
g_index(df = dat2, id = "id", cit = "citations", plot = FALSE)

# Create another example data frame
dat3 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a, b, c", "b, d", "c", "d", "e, g", "f", "g"),
  id = c(123, 234, 345, 456, 567, 678, 789),
  categories = c("a: d: e", "b", "c", "d: g", "e", "f", "g"))

# Calculate g-index and produce plot
g_index(df = dat3, id = "id", cit = "citations", plot = TRUE)
```

h_index

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. Only required when 'plot' parameter is set to "TRUE". Default set to NULL.

<code>cit</code>	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.
<code>plot</code>	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" or "F" to skip plot generation. The default is "FALSE".

Value

h-index value and plot.

Examples

```
# Create an example data frame
dat1 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a; b; c", "b; d", "c", "d", "e; g", "f", "g"),
  id = c("abc123", "bcd234", "def345", "efg456", "fgh567", "ghi678", "hij789"),
  categories = c("a; d; e", "b", "c", "d; g", "e", "f", "g"))

# Calculate h-index
h_index(df = dat1, cit = "citations")

# Create another example data frame
dat2 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a/ b/ c", "b/ d", "c", "d", "e/ g", "f", "g"),
  id = c("123", "234", "345", "456", "567", "678", "789"),
  categories = c("a/ d/ e", "b", "c", "d/ g", "e", "f", "g"))

# Calculate h-index
h_index(df = dat2, id = "id", cit = "citations", plot = FALSE)

# Create another example data frame
dat3 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a, b, c", "b, d", "c", "d", "e, g", "f", "g"),
  id = c(123, 234, 345, 456, 567, 678, 789),
  categories = c("a: d: e", "b", "c", "d: g", "e", "f", "g"))

# Calculate h-index and produce plot
h_index(df = dat3, id = "id", cit = "citations", plot = TRUE)
```

xd_index

xd_index

Description

Calculates 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

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

Arguments

<code>df</code>	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.
<code>cat</code>	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.
<code>id</code>	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.
<code>cit</code>	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.
<code>type</code>	"h" for Hirsch's h-type index or "g" for Egghe's g-type index. Default set to "h".
<code>dlm</code>	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 ";", "/", ":", and ",". The default delimiter is set to ";".
<code>plot</code>	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" or "F" to skip plot generation. The default is "FALSE".

Value

xd-index value and plot.

Examples

```
# Create an example data frame
dat1 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a; b; c", "b; d", "c", "d", "e; g", "f", "g"),
  id = c("abc123", "bcd234", "def345", "efg456", "fgh567", "ghi678", "hij789"),
  categories = c("a; d; e", "b", "c", "d; g", "e", "f", "g"))

# Calculate h-type xd-index
xd_index(df = dat1, cat = "categories", id = "id", cit = "citations")

# Create another example data frame
dat2 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a/ b/ c", "b/ d", "c", "d", "e/ g", "f", "g"),
  id = c("123", "234", "345", "456", "567", "678", "789"),
  categories = c("a/ d/ e", "b", "c", "d/ g", "e", "f", "g"))
```

```
# Calculate g-type xd-index
xd_index(df = dat2, cat = "categories", id = "id", cit = "citations", type = "g", dlm = "/")

# Create another example data frame
dat3 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
                  keywords = c("a, b, c", "b, d", "c", "d", "e, g", "f", "g"),
                  id = c(123, 234, 345, 456, 567, 678, 789),
                  categories = c("a: d: e", "b", "c", "d: g", "e", "f", "g"))

# Calculate h-type xd-index and produce plot
xd_index(df = dat3, cat = "categories", id = "id", cit = "citations", dlm = ":", plot = TRUE)
```

x_index

x_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" for Hirsch's h-type index or "g" for Egghe's g-type index. Default set to "h".
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 ";", "/", ":", and ".". The default delimiter is set to ";".

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" or "F" to skip plot generation. The default is "FALSE".

Value

x-index value and plot.

Examples

```
# Create an example data frame
dat1 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a; b; c", "b; d", "c", "d", "e; g", "f", "g"),
  id = c("abc123", "bcd234", "def345", "efg456", "fgh567", "ghi678", "hij789"),
  categories = c("a; d; e", "b", "c", "d; g", "e", "f", "g"))

# Calculate h-type x-index
x_index(df = dat1, kw = "keywords", id = "id", cit = "citations")

# Create another example data frame
dat2 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a/ b/ c", "b/ d", "c", "d", "e/ g", "f", "g"),
  id = c("123", "234", "345", "456", "567", "678", "789"),
  categories = c("a/ d/ e", "b", "c", "d/ g", "e", "f", "g"))

# Calculate g-type x-index
x_index(df = dat2, kw = "keywords", id = "id", cit = "citations", type = "g", dlm = "/")

# Create another example data frame
dat3 <- data.frame(citations = c(0, 1, 1, 2, 3, 5, 8),
  keywords = c("a, b, c", "b, d", "c", "d", "e, g", "f", "g"),
  id = c(123, 234, 345, 456, 567, 678, 789),
  categories = c("a: d: e", "b", "c", "d: g", "e", "f", "g"))

# Calculate h-type x-index and produce plot
x_index(df = dat3, kw = "keywords", id = "id", cit = "citations", dlm = ",", plot = TRUE)
```

Index

`g_index`, [2](#)

`h_index`, [3](#)

`x_index`, [6](#)

`xd_index`, [4](#)