Obtaining full references

6 Oct 2015

The following example illustrates how to check and find the full references and DOI for the publications used in COMADRE. See also Supporting Information Appendix S4 below.

The code relies on the package `rcrossref` (https://github.com/ropensci/rcrossref) which queries *CrossRef*, an official Digital Object Identifier (DOI) Registration Agency of the International DOI Foundation. Thus you will first need to load the package (and install it if necessary):

```
library(rcrossref)
```

First, set the working directory to where the COMADRE R data object has been saved, and load the data therein:

```
setwd("~/Downloads/COMADRE/")
load("COMADRE_v.1.0.0.RData")
```

The `rcrossref` package has a convenient function, `cr_search_free` which conducts a free-text search of the CrossRef database. To use it, one needs to provide some query text, so in this case we can simply create a text string by concatenating the authors, journal and year of publication from COMADRE.

For example, to obtain the full reference and DOI for the matrices for the koala, *Phascolarctos cinereus*, first identify the pertinent row(s) in the metadata:

```
id <- which(comadre$metadata$SpeciesAccepted == "Phascolarctos_cinereus")
length(id)
```

```
> [1] 5
```

Then use this information to obtain the source information (authors, journal and year of publication) for the 5 matrices:

```
temp <- comadre$metadata[id, c("Authors", "Journal", "YearPublication")]
head(temp)
```

```
>                                                Authors   Journal
> 1272       Baxter; McCarthy; Possingham; Menkhorst; McLean Cons Biol
> 1273 Rhodes; Ng; de Villiers; Preece; McAlpine; Possingham Biol Cons
> 1274 Rhodes; Ng; de Villiers; Preece; McAlpine; Possingham Biol Cons
> 1275 Rhodes; Ng; de Villiers; Preece; McAlpine; Possingham Biol Cons
> 1276 Rhodes; Ng; de Villiers; Preece; McAlpine; Possingham Biol Cons
>      YearPublication
> 1272            2006
> 1273            2011
> 1274            2011
> 1275            2011
> 1276            2011
```

Next, paste these pieces of information together to form a single search string for each matrix. One can optionally ask R to return the unique set of values:

```
x<-apply(temp,1,paste,collapse = " ")
x<-unique(x)
x
```

```
> [1] "Baxter; McCarthy; Possingham; Menkhorst; McLean Cons Biol 2006"
> [2] "Rhodes; Ng; de Villiers; Preece; McAlpine; Possingham Biol Cons 2011"
```

This shows that the data are from 2 publications, for which to obtain source information via CrossRef. The `cr_search_free` function returns a `data.frame`, with a column called `doi` that contains the returned Digital Object Identifiers for the publications.

```
temp <- cr_search_free(x)
temp$doi
```

```
> [1] "http://dx.doi.org/10.1111/j.1523-1739.2006.00378.x"
> [2] "http://dx.doi.org/10.1371/journal.pone.0092430"
```

Armed with the DOI, it is easy to obtain the full title, author list etc. from CrossRef in a range of formats using the function `cr_cn`. This uses the raw DOI, without the `http://dx.doi.org/` prefix. Therefore this prefix must first be stripped from the query using `gsub`.

```
doiValues <- gsub("http://dx.doi.org/","",temp$doi)
cr_cn(dois=doiValues,format = "text", style = "apa")
```

```
> [[1]]
> [1] "BAXTER, P. W. J., McCARTHY, M. A., POSSINGHAM, H. P., MENKHORST, P. W., & McLEAN, N. (2006).
>
> [[2]]
> [1] "Ng, C. F., Possingham, H. P., McAlpine, C. A., de Villiers, D. L., Preece, H. J., & Rhodes, J
```