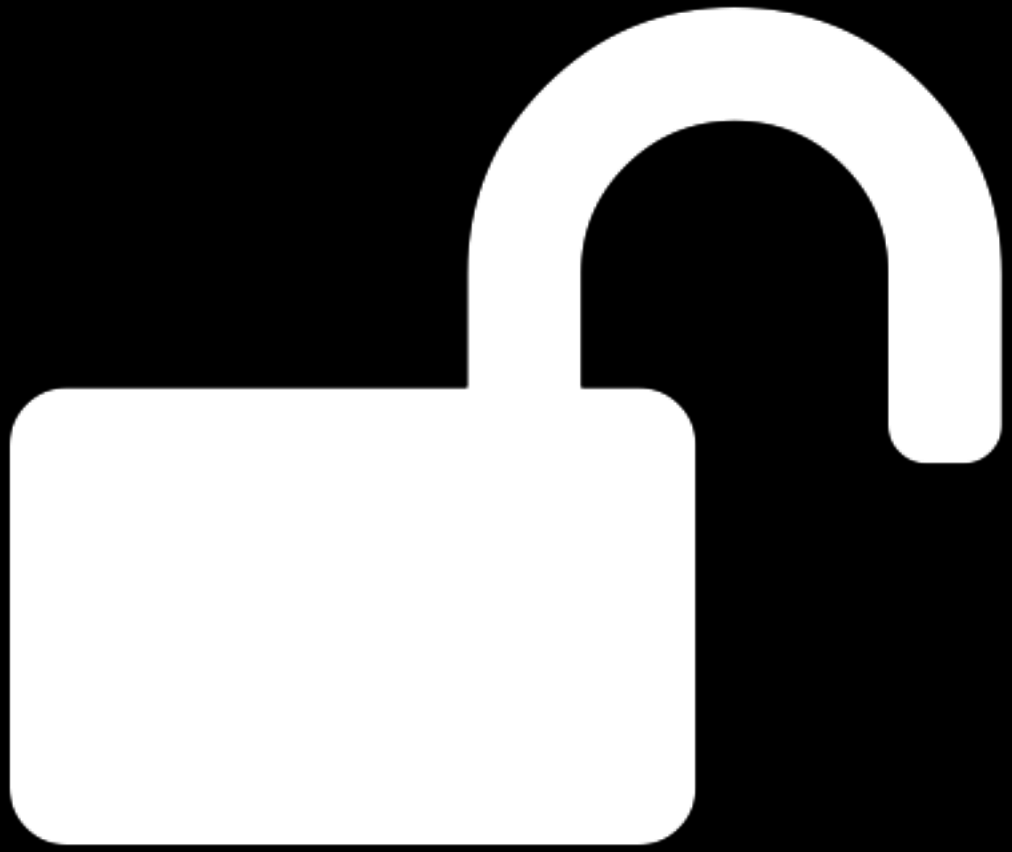
R-based tools for open and collaborative science

@recology_
Scott Chamberlain

Science needs to be more open



We build on the knowledge of others

http://everyoneknowsbest.files.wordpress.com/2008/08/bodysculpture.jpg

Less mistakes

More things can happen b/c data is open

The public paid for it!

http://everyoneknowsbest.files.wordpress.com/2008/08/bodysculpture.jpg

But we need tools to do it!!!!!

http://www.fotopedia.com/items/flickr-4796633039

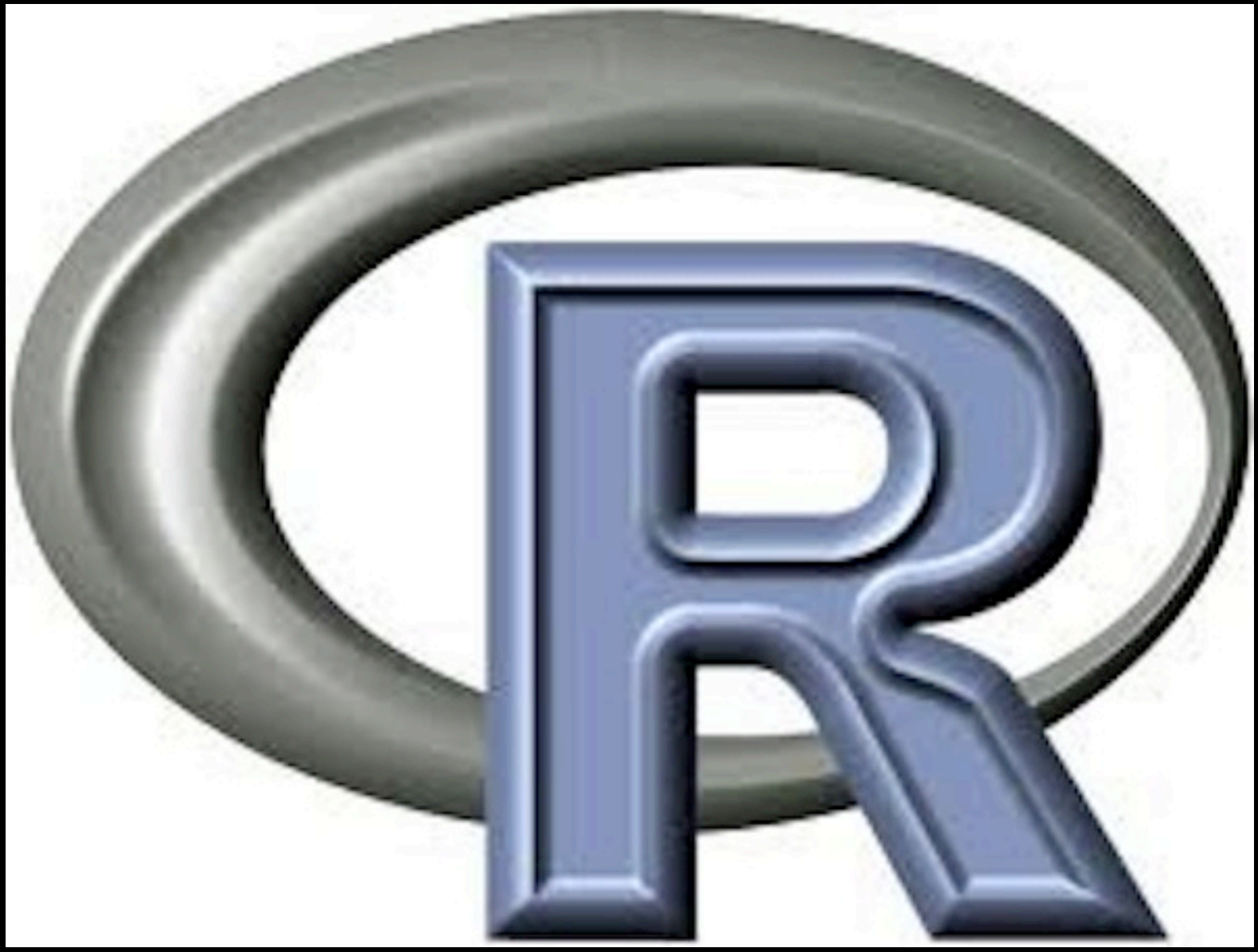
What kinds of tools? Not these

These!!!!!!

What does an ecologist do?

- Collect data
- Manipulate data
- Visualize
- Analyze
- Write

R is a good solution



Why?

- R is Open source = Free + Rapid change
- R = entire workflow in 1 place
- R = reproducible science

Get some data from the web

```
library(RCurl); library(RJSONIO)
dat <- fromJSON(getURL("https://api.github.com/users/hadley/repos"))
```

Manipulate the data

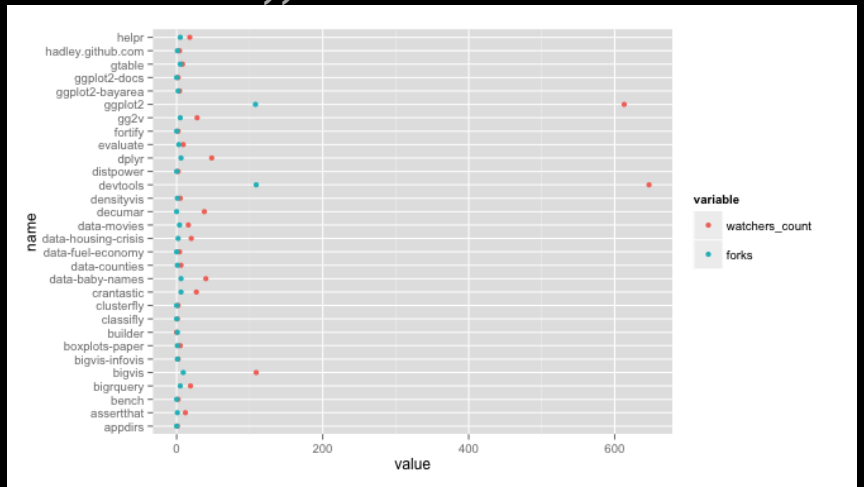
```
library(plyr); library(reshape2)
dat_melt <- melt(ldply(dat, function(x) data.frame(x[names(x) %in%
                    c("name","watchers_count","forks")])))
```

Run some statistical model

```
lm(value ~ variable, data = dat_melt)
```

Visualize results

```
library(ggplot2)
ggplot(dat_melt, aes(name, value, colour = variable)) +
        geom_point() +
        coord_flip()
```



Write the paper

```
# Introduction...
```

Data increasingly on the web

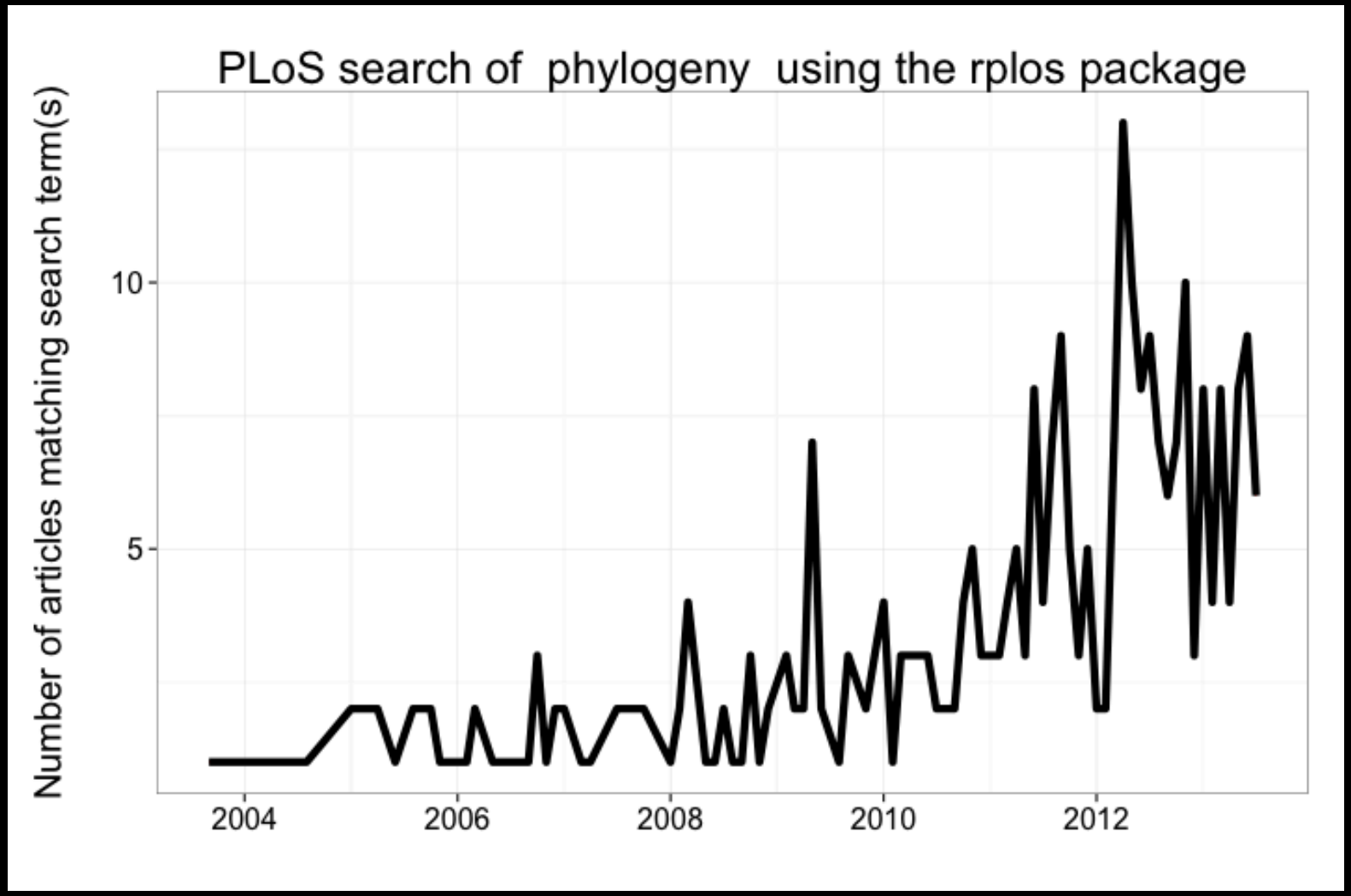
ITIS

GBIF

The toolbelt

Literature

```
library(rplos)
plot_throughtime('phylogeny', 300) + geom_line(size=2)
```



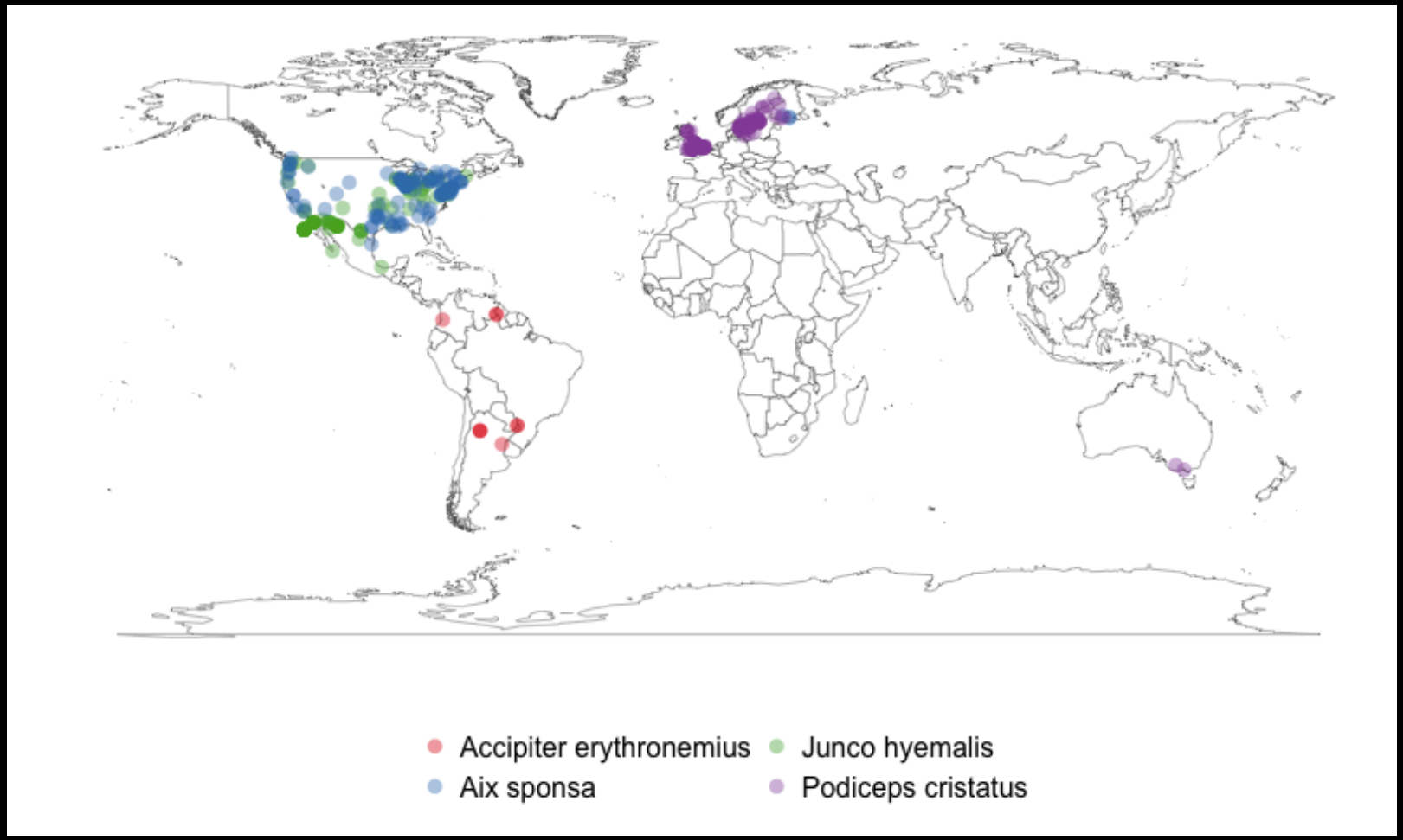
Taxonomy

```
library(taxize)
classification("Abies procera",  db = "itis")
```

rankName	taxonName	tsn
Kingdom	Plantae	202422
Subkingdom	Viridaeplantae	846492
Infrakingdom	Streptophyta	846494
Division	Tracheophyta	846496
Subdivision	Spermatophytina	846504
Infradivision	Gymnospermae	846506
Class	Pinopsida	500009
Order	Pinales	500028
Family	Pinaceae	18030
Genus	Abies	18031
Species	Abies procera	181835

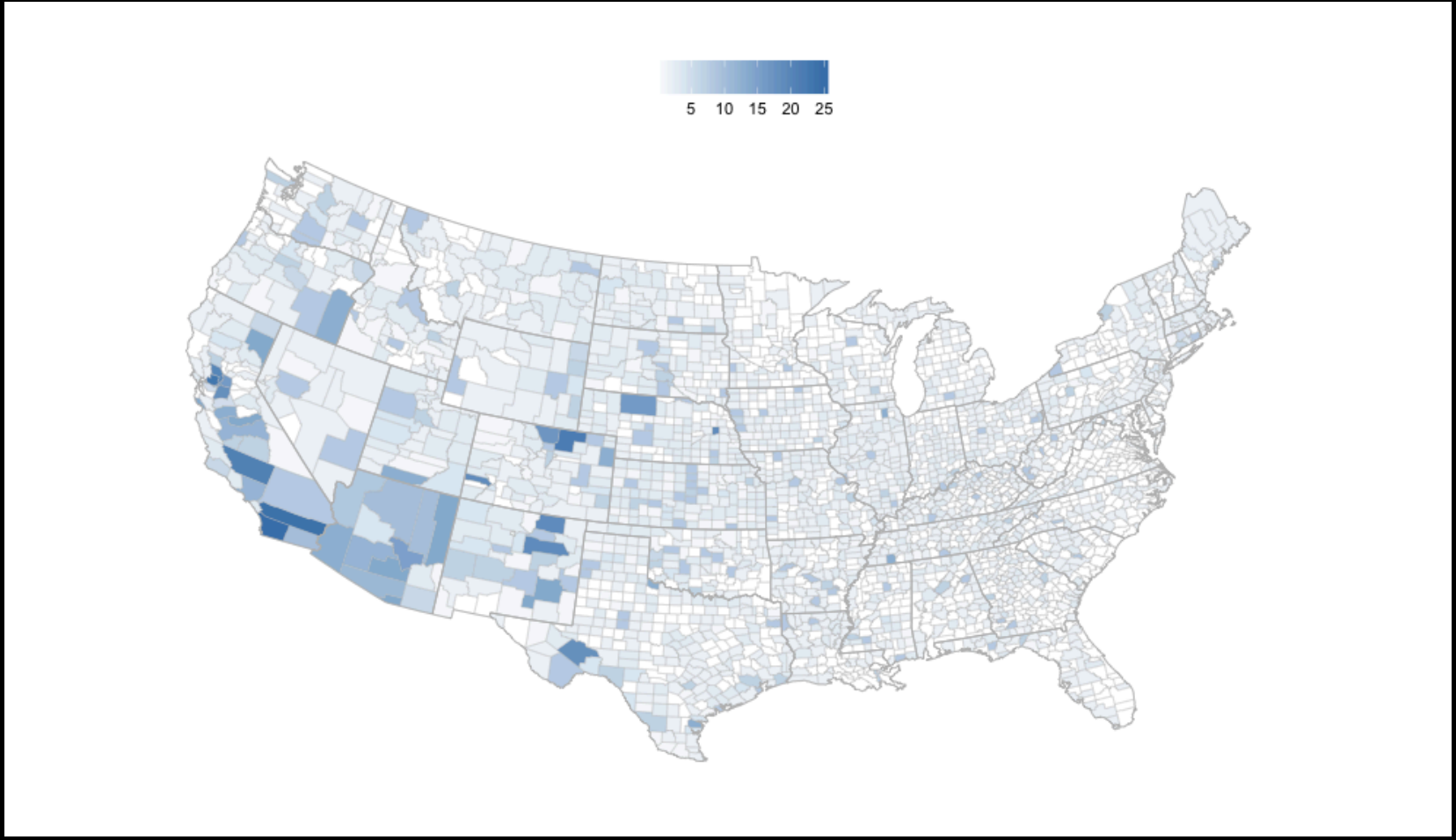
Species occurrences from GBIF

```
library(rgbif)
splist <- c('Accipiter erythronemius', 'Junco hyemalis', 'Aix sponsa', 'Podiceps cristatus')
out <- occurrencelist_many(splist)
gbifmap_list(out)
```



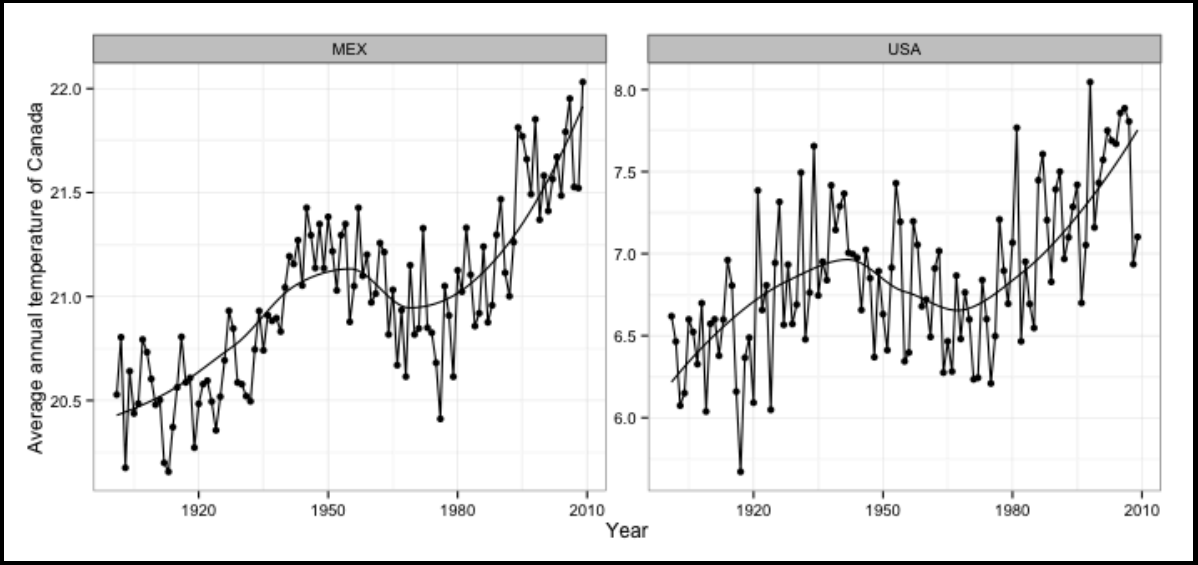
Occurrence from USGS's BISON service

```
library(rbison)
out <- bison(species="Helianthus annuus", count=500)
bisonmap(input=out, tomap="county")
```



Climate data from the World Bank

```
library(rWBclimate)
country.list <- c("USA", "MEX")
country.dat <- get_historical_temp(country.list, "year")
ggplot(country.dat, aes(x = year, y = data, group = locator)) +
  geom_point() +
  geom_path() +
  labs(y="Average annual temperature of Canada", x="Year") +
  theme_bw() +
  stat_smooth(se = F, colour = "black") +
  facet_wrap(~locator, scale = "free")
```



Take action!

FontAwesome http://fortawesome.github.io/Font-Awesome/
fontawesome 2 png https://github.com/odyniec/font-awesome-to-png
Presentation available here: http://bit.ly/16tuVbu