R output to MS Word and HTML

Aniko Szabo

Medical College of Wisconsin
Milwaukee, WI

June 2, 2011

Outline

1. `R2wd`: Creating MS Word files using R

2. `R2HTML`: Creating HTML files using R
   - Console-like output
   - Explicit commands
   - HTML file with R chunks

3. `R2wd` vs `R2HTML`: pros and cons

R2wd: Creating MS Word files using R

R to Word:R2wd

- Uses the StatconnDCOM server to communicate with Word
- Windows only
- Works only on 32-bit version of R (>=2.12.0) and MS Word

Installing R2wd

- R2wd is on CRAN
- At your first attempt to load it, an informational message tells you to use `installstatconnDCOM()` to install the StatconnDCOM server
- This will download the required files from http://rcom.univie.ac.at/. You will have to step through a typical Windows installation process

Using R2wd

- `wdGet()` opens a new Word document, or establishes connection to one that is an already open
- The R script should contain explicit commands to put titles, text, plots, tables, etc into the Word document:
  - `wdTitle`,`wdHeading`: insert title/headings
  - `wdWrite`,`wdBody`,`wdNormal`,`wdVerbatim`: insert text
  - `wdSection`, `wdSubsection`: insert section breaks + bookmark + heading
  - `wdPlot`, `wdTable`: insert plot/table
- `wdSave(`*`filename`*`)` saves the file, `wdQuit()` cuts the connection

R2wd example

```
library(R2wd)

wdGet()
wdTitle("Simple example of using R2wd")

wdSection("Introduction")

data(iris)
wdBody(c("Consider Fisher's famous iris data set.",
 "It has sepal and petal length and width of",
 sprintf("a total of %d flowers from %d species.",
 nrow(iris), nlevels(iris$Species))))

wdBody("The first few lines of the data are shown in Table
    1.")
wdTable(head(iris), caption="Data for the first 6 flowers.")
...
```

R2wd example

```
...
wdSection("Analyzing petal widths")
wdBody(c("Figure 1 shows a boxplot of petal widths by
    species",
   "- there is a large apparent separation of this
       measurement."))
wdPlot(Petal.Width ~ Species, data=iris, plotfun=boxplot,
       width=6, height=4)

wdBody("The separation can be formally confirmed via ANOVA:"
    )

aov1 <- aov(Petal.Width ~ Species, data=iris)
wdVerbatim(capture.output(summary(aov1)))

wdSave("R2wdExample.docx")
wdQuit()
```

The resulting Word file

R2HTML: Creating HTML files using R

R to HTML: `R2HTML`

- Uses plain text output via `cat` to create HTML files
- Works with any operating system
- Available on CRAN – no special installation is needed
- Three usage modes:
  1. a record of the console output
  2. explicit commands from an R script (~`R2wd`)
  3. HTML document with R chunks (~`Sweave`)

Using R2HTML: recording console output

- Start with `HTMLStart()`
- End with `HTMLStop()`
- Creates a "framed" web-page with the executed commands in the left frame linked to the output in the right frame
- Use `HTMLPlot()` to output contents of current graphics window as a plot

R2HTML console output example

```
library(R2HTML)
HTMLStart(outdir=getwd(),file="R2htmlExample1",echo=TRUE)
data(iris)
head(iris)

boxplot(Petal.Width ~ Species, data=iris)
HTMLPlot()

aov1 <- aov(Petal.Width ~ Species, data=iris)
summary(aov1)

HTMLStop()
```

The resulting HTML file

Using R2HTML: explicit commands

- Place code between `HTMLInitFile` and `HTMLEndFile`
- Use `HTML` to output to file – has over 150 methods!
- Some special commands: `HTML.title`, `HTMLhr`, `HTML.latex` – AsciiMathML

Example: R2HTML explicit commands

```
HTMLInitFile(outdir=getwd(), filename="R2htmlExample2")

HTML.title("Simple example of using R2HTML")
HTML.title("Introduction", HR=3)

HTML(sprintf("Consider Fisher's famous iris data set.
   It has sepal and petal length and width of
   a total of %d flowers from %d species.",
   nrow(iris), nlevels(iris$Species)))

HTML("The first few lines of the data are shown in Table 1."
    )
HTML(head(iris), caption="Data for the first 6 flowers.")
...
```

Example: R2HTML explicit commands

```
HTML.title("Analyzing petal widths", HR=3)
HTML("Figure 1 shows a boxplot of petal widths by species -
   there is a large apparent separation of this measurement.
      ")
boxplot(Petal.Width ~ Species, data=iris)
HTMLplot(Caption="Variation of petal width by species",
        Width=600, Height=400)

HTML("The separation can be formally confirmed via ANOVA:")
aov1 <- aov(Petal.Width ~ Species, data=iris)
HTML(summary(aov1))

HTMLEndFile()
```

The resulting HTML file

Using R2HTML: HTML file with R chunks

- R code chunks are inserted into an HTML file using noweb syntax: <<>>=
- Inline expression can be included with <Sexpr ...>
- File processed through *Sweave*:

```
Sweave("R2htmlExample3.snw",driver=RweaveHTML(),
       cssfile="R2HTML.css")
```

Example: HTML file with R chunks

```
<html>
<head><h1>
Simple example of using R2HTML via Sweave
</h1></head>

<h3> Introduction </h3>

<<Intro,echo=false>>=
data(iris)
@

<p>Consider Fisher's famous iris data set.  It has sepal and
petal length and width of a total of <Sexpr nrow(iris)>
flowers from <Sexpr nlevels(iris$Species)> species </p>

<p>The first few lines of the data are shown in Table 1.</p>

<<IrisHead, results=html, echo=false>>=
HTML(head(iris), caption="Data for the first 6 flowers.")
...
@
```

Example: HTML file with R chunks

```
...
<h3>Analyzing petal widths</h3>
<p>Figure 1 shows a boxplot of petal widths by species -
   there is a large apparent separation of this measurement.
       </p>

<<IrisPlot, echo=false, fig=true,
   caption=Variation of petal width by species,
   width=600, height=300, HTMLwidth=600, HTMLheight=300>>=
boxplot(Petal.Width ~ Species, data=iris)
@

<p>The separation can be formally confirmed via ANOVA:</p>

<<IrisAnova, results=html, echo=false>>=
  aov1 <- aov(Petal.Width ~ Species, data=iris)
  HTML(summary(aov1))
@
</html>
```

The resulting HTML file

R2wd vs R2HTML: pros and cons

R2wd vs R2HTML

R2wd

- \+ Produces document in a commonly used/requested format
- − Only available on Windows
- − Not well developed yet; no functions/packages for complicated formatting

R2HTML

- \+ Works on all platforms, no proprietary formats
- \+ Better developed: can deal with lots of data types; many packages (e.g. `xtable`) can create nicely formatted output
- \+ Multiple usage modes provide flexibility
- − Result cannot be readily edited (without HTML knowledge)
- − If plots are used, results in many files