

Pagedown for PDF publications with HTML and CSS

R Addicts Meetup

Christophe Dervieux

April 18, 2019

About

Short introduction

User of  for 7 years

 RTE (Electricity Transmission Network)¹

Admin R², Developer.

In support of projects and users.



@cderv



@chrisderv

[1] this presentation is not in an official setting - just a personal sharing.

[2] <https://rviews.rstudio.com/2017/06/21/analytics-administration-for-r/>

Rmarkdown to produce PDF ... but without Latex

A crazy idea?

Bookdown contest submission: a pulp magazine template (a demo with CSS paged media) 

[R Markdown](#) [bookdown](#) [bookdown-contest](#)



rlesur

sept. '18

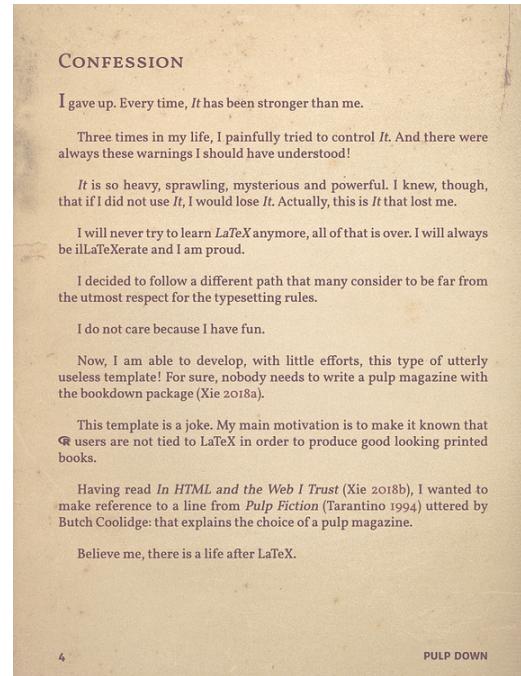
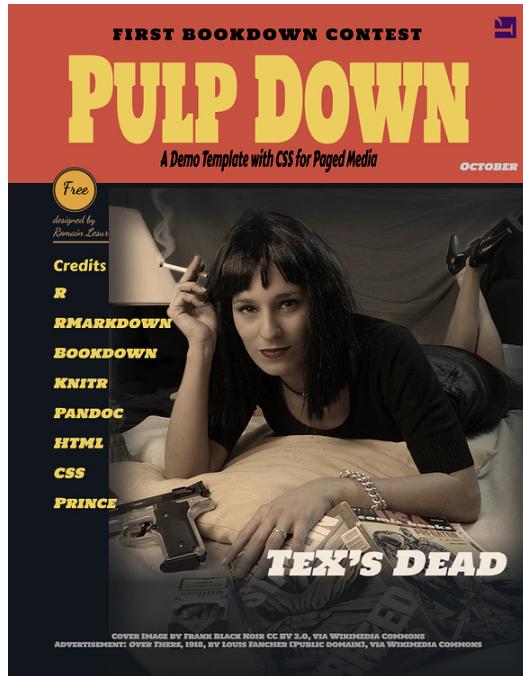
The usual way to convert a bookdown to PDF relies on LaTeX.
In order to customize the rendered document, one has to pass the painful LaTeX learning curve.

However, there is an underrated way to produce a PDF with HTML and CSS only.
This template was developed in order to introduce this alternative.

A PDF magazine with Rmarkdown and HTML / CSS?

Rmarkdown to produce PDF ... but without Latex

It's possible !



- Demo Project : **Ronsur / pultdown**
- Using the **Prince XML** Software Here



pagedown

Paginated html with Rmarkdown

Following the competition, a collaboration Yihui Xie and Romain Lesur

Allow the creation of documents **layout for publication pdf**
in HTML and CSS

Version 0.2 on the CRAN  Experimental

Prefer an installation from  for the latest news and corrections

```
remotes::install_github('rstudio/pagedown')
```

Requires **Pandoc> 2** - includes in RStudio> 1.2 (Release early April)

To see the result: Google Chrome or **Chromium**

use

Many templates available!

| File -> New File -> R Markdown -> From Template

Template for template

Paged HTML document (`pagedown::html_paged`)

<https://pagedown.rbind.io/>

The screenshot shows a PDF document titled "pagedown_Creat..." with a table of contents. The title page contains the book's title, authors, and date. The table of contents lists various sections and their page numbers.

pagedown: Create Paged HTML Documents for Printing from R Markdown

A Less Traveled Road to PDF and Printing

Yihui Xie and Romain Lesur

2018-10-29

Contents

1 Introduction	5
1.1 CSS for print	5
1.2 CSS frameworks	5
2 Paged.js	7
3 Applications	9
4 Supported features	11
4.1 Lists of Tables and Figures	11
4.2 Links	11
4.2.1 Automatic links	11
4.2.2 Inline links	12
4.3 Tests	13
4.3.1 A list	13
4.3.2 Test MathJax	17
4.3.3 Miscellaneous tests used during pagedown development	17
Bibliography	21

Template for template

Journal of Statistical Software article (pagedown::jss_paged>)

The screenshot shows the RStudio interface with a code editor on the left and a viewer pane on the right.

Code Editor (JSS.Rmd):

```
1 ---  
2 author:  
3   - name: Yihui Xie  
4     affiliation: RStudio  
5     address: |  
6       First line  
7       Second line  
8     email: <name@company.com>  
9     url: <http://rstudio.com>  
10    - name: Romain Lesur  
11 title:  
12   [HTML]{.proglang} and [CSS Paged Media]{.proglang} for  
Scientific Publications with the [pagedown]{.pkg}  
[R]{.proglang} Package  
shorttitle: "[pagedown]{.pkg}: Paginate the [HTML]{.proglang}  
Output of [R Markdown]{.pkg} with [CSS]{.proglang} for Print"  
14 abstract:  
15   Use the paged media properties in [CSS]{.proglang} and the  
[JavaScript]{.proglang} library [Paged.js]{.pkg} to split the  
content of an [HTML]{.proglang} document into discrete pages.  
Each page can have its page size, page numbers, margin boxes,  
and running headers, etc. Applications of this package  
include books, letters, reports, papers, business cards,  
resumes, and posters.  
16 # at least one keyword must be supplied  
17 keywords:  
18   - "[HTML]{.proglang}"  
19   - "[CSS]{.proglang}"  
20   - "[markdown]{.proglang}"  
21   - scientific publication  
22   - literate programming  
23 links-to-footnotes: true  
24 paged-footnotes: true  
25 link-citations: true  
26 bibliography: index.bib  
27 output:  
28   pagedown::jss_paged:  
29     self_contained: false  
30 ---  
31 # Introduction  
33  
34 When talking about PDF and printing, we often think of tools  
like  $\$LaTeX$  and Microsoft Word. When talking about  
[HTML]{.proglang} and [CSS]{.proglang}, we may have never  
imagined their possible off-screen use such as printing to  
PDF.  
35  
36 Can we print a book with [HTML]{.proglang} and  
[CSS]{.proglang}? W3C published the first working  
draft1 (https://www.w3.org/1999/06/WD-css3-page-19990623) on  
"Paged Media Properties for CSS(3)". which was
```

Viewer Pane:

The viewer pane displays the rendered *Journal of Statistical Software* article. It includes a header with the journal title, volume, issue, and DOI, followed by a large image, the abstract, and the introduction section.

Abstract:

Use the paged media properties in CSS and the JavaScript library Paged.js to split the content of an HTML document into discrete pages. Each page can have its page size, page numbers, margin boxes, and running headers, etc. Applications of this package include books, letters, reports, papers, business cards, resumes, and posters.

Keywords: HTML, CSS, markdown, scientific publication, literate programming.

1. Introduction

When talking about PDF and printing, we often think of tools like $\$LaTeX$ and Microsoft Word. When talking about HTML and CSS, we may have never imagined their possible off-screen use such as printing to PDF.

Can we print a book with HTML and CSS? W3C published the first working draft¹ on "Paged Media Properties for CSS(3)", which was last updated in 2013. Although the working draft has been there for nearly two decades, it is still not common to see authors write or print books with HTML and CSS. The main reason is that the W3C specs are still in the draft mode, so most web browsers

Template for template

A CV !(pagedown::html_resume)

Lijia Yu's resume.pdf (1 page) ▾

LIJIA YU
CURRENTLY SEARCHING FOR A PHD STUDENT POSITION

EDUCATION

2010 • **Beijing University of Chemical Technology**
B.S. in Information and Computing Sciences
Beijing, China

2014 • **University of Chinese Academy of Sciences**
M.S. in Bioinformatics
Beijing, China

RESEARCH EXPERIENCE

2011 | 2014 • **Graduate Research Assistant**
Beijing Institute of Genomics, Chinese Academy of Sciences
• Performed computational biology research towards understanding regulation of alternative splicing in human and mouse transcriptome.
• Found EGFR pathway related mutations, aimed to understand the impacts of cancer mutations on EGFR signaling pathway.
Beijing, China

2015 | 2016 • **Bioinformatician**
My Health Gene Technology Co., Ltd.
• Investigated how cancer cells spread to other parts of the body at the single cell level.
Beijing, China

2016 | 2018 • **Visiting Scientist**
University of Alabama at Birmingham
• Investigated the role of mitochondria in development of cancer.
• Investigated the evolution of genome architecture and its role in important evolutionary events.
• Detected thrombotic thrombocytopenic purpura related mutations in multiple patients' blood genome.
AL, USA

CONTACT INFO

✉ lijia.yu@outlook.com
⌚ github.com/yulijia
📞 +1 000-000-0000
For more information,
please contact me via
email.

SKILLS

Experienced in statistical analysis,
statistical learning models,
and optimization methods.

Full experience with next generation sequencing data analysis.

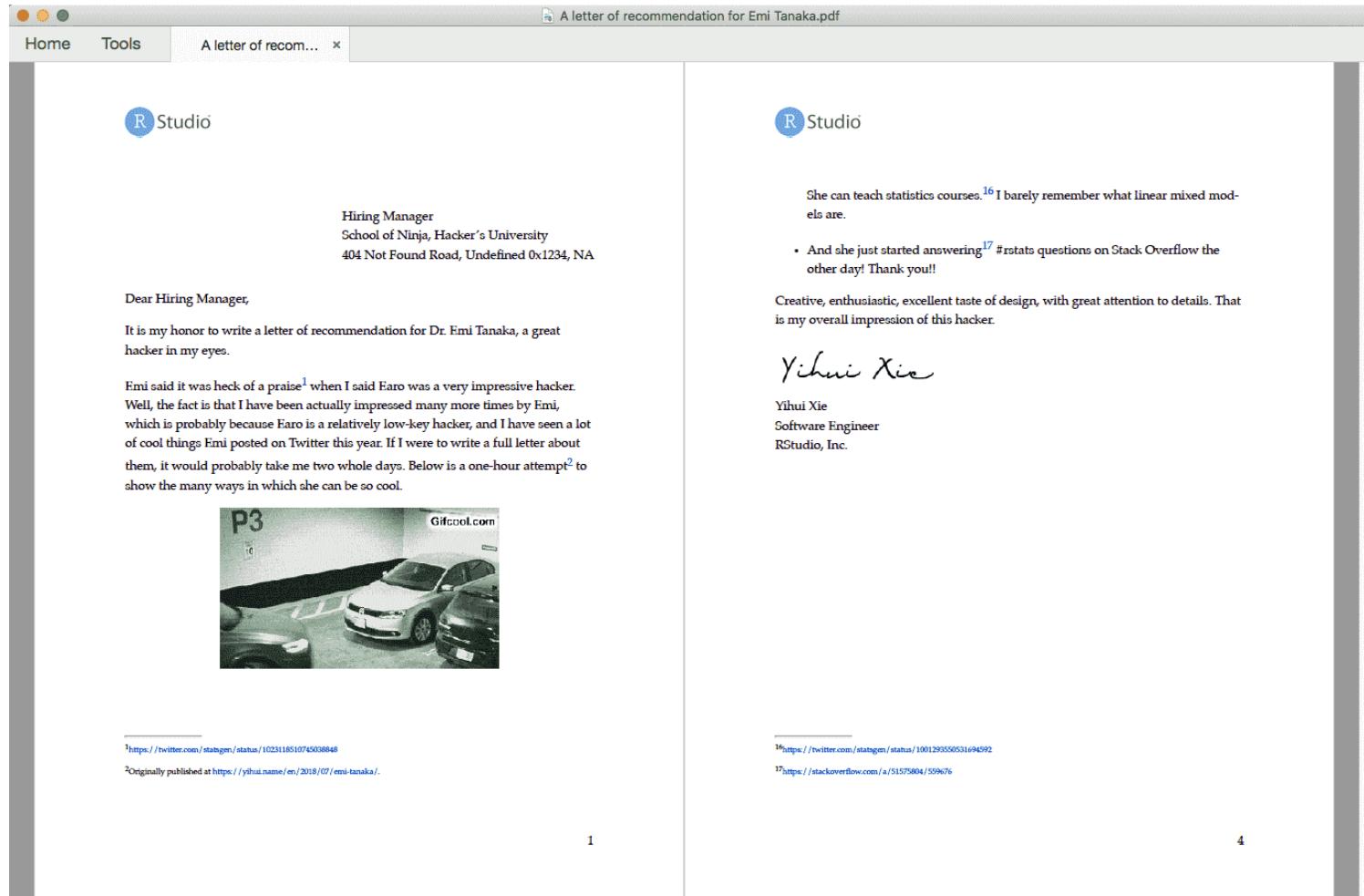
Highly skilled in R, Bash, Perl,
Python, LaTeX

PROFESSIONAL EXPERIENCE

2014 • **Data Scientist, intern**

Template for template

A letter (pagedown::html_letter)



The screenshot shows an RStudio interface with a PDF document titled "A letter of recommendation for Emi Tanaka.pdf". The document contains the following content:

R Studio

Hiring Manager
School of Ninja, Hacker's University
404 Not Found Road, Undefined 0x1234, NA

Dear Hiring Manager,

It is my honor to write a letter of recommendation for Dr. Emi Tanaka, a great hacker in my eyes.

Emi said it was heck of a praise¹ when I said Faro was a very impressive hacker. Well, the fact is that I have been actually impressed many more times by Emi, which is probably because Faro is a relatively low-key hacker, and I have seen a lot of cool things Emi posted on Twitter this year. If I were to write a full letter about them, it would probably take me two whole days. Below is a one-hour attempt² to show the many ways in which she can be so cool.



¹<https://twitter.com/statgen/status/1023118510745038848>

²Originally published at <https://yihui.name/en/2018/07/emi-tanaka/>.

R Studio

She can teach statistics courses.¹⁶ I barely remember what linear mixed models are.

- And she just started answering¹⁷ #rstats questions on Stack Overflow the other day! Thank you!!

Creative, enthusiastic, excellent taste of design, with great attention to details. That is my overall impression of this hacker.

Yihui Xie

Yihui Xie
Software Engineer
RStudio, Inc.

¹⁶<https://twitter.com/statgen/status/100129550531694992>

¹⁷<https://stackoverflow.com/a/51575804/59676>

Template for template

Poster on the theme of Relaxed (pagedown::poster_relaxed)

<https://pagedown.rbind.io/poster-relaxed>

Full HTML / CSS here - formatting is done without JavaScript

Template for template

Book "at the bookdown" (`pagedown::book_crc`)

Yihui Xie

bookdown: Authoring Books and Technical Documents with R Markdown

Contents

List of Tables	vii
List of Figures	ix
Preface	xi
About the Author	xix
1 Introduction	1
1.1 Motivation	2
1.2 Get started	3
1.3 Usage	4
1.4 Two rendering approaches	8
1.5 Some tips	9
2 Components	11
2.1 Markdown syntax	11
2.1.1 Inline formatting	11
2.1.2 Block-level elements	12
2.1.3 Math expressions	14
2.2 Markdown extensions by bookdown	15
2.2.1 Number and reference equations	15
2.2.2 Theorems and proofs	17
2.2.3 Special headers	22
2.2.4 Text references	23
2.3 R code	24
2.4 Figures	25

iii

Examples in nature?

Examples in nature?

How to quickly
produce statisti-
cal reports?

The UNHCR "*R-Cookbook*"

Version: 2019-04-08

Rendering possible by Paged.js

An open-source library for
paging content in a browser

Paged.js is a Polyfill

A polyfill is simple. It is a set of **functions to simulate** ,
on an old web browser, **features** that are **not natively**
available .¹

Here, *Paged.js* comes to make available W3C specifications on **Paged Media** and **Generated Content for Paged Media** on recent browsers. So it comes complete Chrome that **supports** only a small part.

! A JS library in full development

(1) <https://blog.eleven-labs.com/fr/tutoriel-polyfill/>

The principles

CSS Paged Media & paged.js

The CSS standard for Paged Media

CSS for Paged Media Standard is a piece of the CSS W3C standard:

- CSS Paged Media Module Level 3
- CSS Generated Content for Paged Media Module
- CSS Page Floats
- CSS Fragmentation Module Level 3

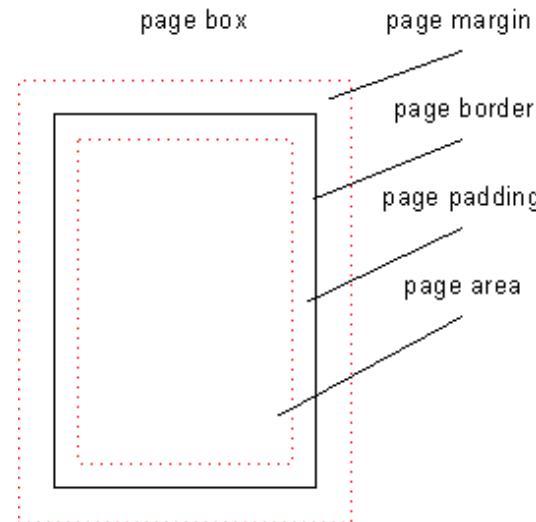
Used in Publishing by Publishers (
Streamlining CSS Print Design with Sass)

The principles

CSS Paged Media & paged.js

The CSS
standard for
Paged Media

The Page
template



```
@page {  
    size: A5;  
}  
@page {  
    margin: 1in 2in .5in 2in;  
}
```

The principles

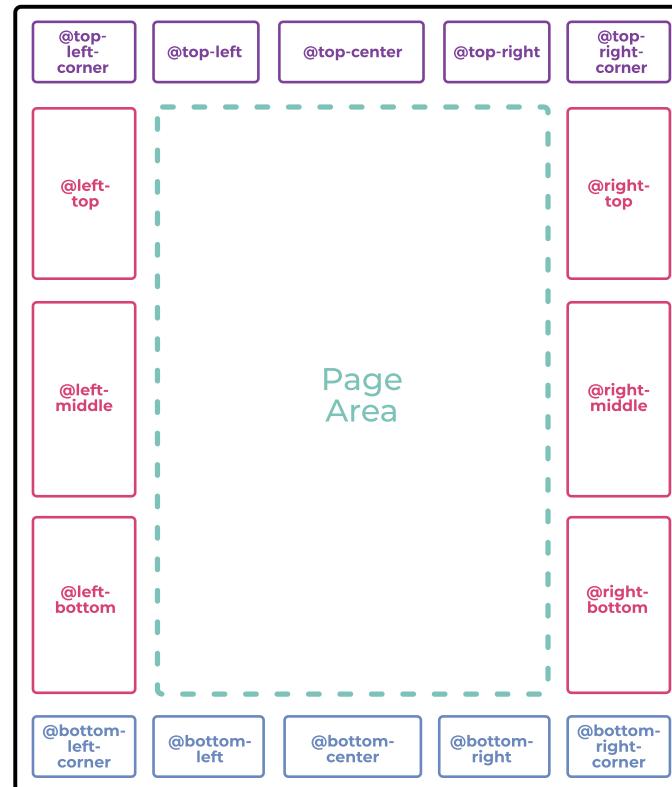
CSS Paged Media & paged.js

The CSS
standard for
Paged Media

The Page
template

The "Margin
Box"

- Cutting a page and margins in identifiable boxes



- In each box, classic CSS rules

The principles

CSS Paged Media & paged.js

The CSS
standard for
Paged Media

The Page
template

The "Margin
Box"

```
@page {  
    size: a4;  
    margin: 15mm;  
    @top-center {  
        content: "Mon Document Paginé";  
    }  
    @bottom-right-corner {  
        content: "p. " counter(page);  
    }  
}  
  
@page :right {  
    @bottom-right-corner {  
        content: counter(page) "/" counter(pages);  
    }  
}  
  
@page :left {  
    @bottom-left-corner {  
        content: counter(page) "/" counter(pages);  
    }  
}
```

The principles

CSS Paged Media & paged.js

The CSS
standard for
Paged Media

The Page
template

The "Margin
Box"

Leaping
pages

- Page break before all titles 1

```
h1 {  
    break-before: page;  
}
```

- Avoid jumps in a table

```
table {  
    break-inside: avoid;  
}
```

How to produce PDF?

CSS paged media via pagedown

Use
pagedown
templates

```
---
```

```
title: "pagedown: Create Paged HTML"
shorttitle: "Printing R Markdown with CSS"
author: "Yihui Xie and Romain Lesur"
date: '2019-04-18'
subtitle: A Less Traveled Road to PDF
output:
  pagedown::html_paged:
    toc: true
    self_contained: false
toc-title: Contents
paged-footnotes: true
lot: true
lof: true
bibliography: index.bib
---
```

Consult the documentation for the different options of each format

How to produce PDF?

CSS paged media via pagedown

Use
pagedown
templates

Preview the
document

Since HTML rendering uses Javascript and CSS, you need to preview with a local or remote web server.¹

Since R, the "magic" of xaringan to help us:

```
# prévisualiser avec un server local  
  
# le document Rmd  
xaringan::inf_mr("your-document.Rmd")  
  
# le document actif dans RStudio  
xaringan::inf_mr()
```

or use the Addin **Infinite Moon Reader**

The RStudio Viewer also works because it uses a local server.

How to produce PDF?

CSS paged media via pagedown

Use
pagedown
templates

Preview the
document

PDF printing

The simplest: Use the Print PDF feature of Chrome when the page is loaded.

The most automated: A special feature `pagedown::chrome_print` to launch chrome printing from R¹:

```
# le html
pagedown::chrome_print("output.html")
# le Rmd qui sera compilé avant impression
pagedown::chrome_print("mondocument.Rmd")
```

This function may or may not work.
Automated printing is complicated and needs to be improved. (taking into account all htmlwidget, javascript interactions, print delays, ...)

(1) To go further in the control of chromium since R, go see `crrri`, package in development to interface chrome and R.

How to produce PDF?

CSS paged media via pagedown

Use
pagedown
templates
Preview the
document

PDF printing

It can also be used for other types of printing:

- Screenshot of a whole page (`format = 'png'`)

```
pagedown::chrome_print(  
  'https://pagedown.rbind.io/poster-relaxed',  
  format = 'png')
```

The screenshot shows a presentation poster with the following sections:

- Motivation**: A section with a red eye icon and a list of bullet points.
- Authors**: A section featuring four authors with their names: Xaringan, R, Hadley, and Jeroen.
- Schedule**: A section with a signature icon and a list of events.
- Content**: A large section containing a mathematical formula $X = \sum_{i=1}^n Y_i$.
- Review and review**: A section with a list of bullet points.
- Cooperating**: A section with a list of bullet points.
- Publishing**: A section with a list of bullet points.
- Bibliography**: A section with a list of references.

How to produce PDF?

CSS paged media via pagedown

Use
pagedown
templates

Preview the
document

PDF printing

It can also be used for other types of printing:

- Screenshot of a page element (`selector = '#authors'`)

```
pagedown::chrome_print(  
  'https://pagedown.rbind.io/poster-relaxed',  
  format = 'png',  
  selector = '#authors' )
```

The screenshot shows a PDF page with a red header bar containing the word "Authors". Below the header, there is a paragraph of text: "We are a team of shinobi and kuniochi who wish to share the fun and secrets of the `xaringan` package with you." To the right of the text are four author profiles, each with a small photo and a caption:

- Emi Tanaka**: Lead author and the ninja theme author. Emi laid out the first sketch of the book, and we all believe that the book had been half done.
- Joseph Carrille**: Coauthor of `xaringan`. "Count me in!" replied Joseph when Yihui asked who wanted to co-author the book.
- Eric Nantz**: Host of the R Podcat. Yihui is super to know how much Eric's writing is better than his magnetic voice.
- Yihui Xie**: Author of `xaringan`. Yihui knows a bit about R/HTML/CSS/JS/JSX and he has been a great help because a corner stone of R Markdown.

At the bottom of the page, a note states: "All pictures above are from the authors' Github profiles. This post was created via the R package `pagedown`".

How to produce PDF?

CSS paged media via pagedown

Use
pagedown
templates

Preview the
document

PDF printing

Customize

CSS, CSS, CSS !!

- Edit CSS rules with a new document

```
output:  
pagedown::html_paged:  
  css: [ "my-fonts.css", "default-page", "defaul
```

- Edit by pages and by margins

Edit or Add templates

- Jedi Master of CSS!
- HTML template pandoc: A subject in its own right!

Volunteers? Get inspired by examples in Pagedown

Thank you !

<https://cderv.github.io/presentations/meetupr-pagedown/>