Vivliostyle

Open Source, Web Browser based CSS Typesetting Engine

Shinyu Murakami
Johannes Wilm

Vivliostyle Inc.
http://www.vivliostyle.com

Typical input/output requirements for publishing

Input:

- Word processor
- Web-based editor

Output:

- Ebook
- Web
- Print

Formats used in workflow

Input:

- Word processor: DOC/DOCX
- Web-based editor: (X)HTML

Output:

- Ebook: EPUB (XHTML)
- Web: (X)HTML
- Print: PDF

Typical workflows

Some examples:

- DOC → XML → XHTML/PDF
- DOC → LaTeX → PDF (but what about HTML/EPUB?)
- HTML → XML → XHTML/PDF
- HTML → XML → PDF → HTML

Common issues with conversion:

- Labor intensive
- Every conversion is likely causing some problems. More conversions = more problems

One solution: Switch to XHTML as main format for long text content

Input:

- Word processor: DOCX/DOC
- Web-based editor: XHTML

Output:

- Ebook: XHTML
- Web: XHTML
- Print: XHTML

Style XHTML for output

Cascading Stylesheets can format XHTML for different output

- Epub/Web: Existing CSS rules (first release 1996)
- Print:

  Additionally three CSS specs:

  - CSS Paged media
  - Generated Content for Paged Media
  - CSS Page Floats

Why has XHTML/CSS for print not had its breakthrough so far?

There are existing solutions for XHTML/CSS for print:

- PDFreactor
- PrinceXML
- Antenna House Formatter
- Pagination.JS
- Simplepagination.JS

Renderer-from-scratch solutions

In this category:

- PDFreactor
- PrinceXML
- Antenna House Formatter

Issues:

- High development costs
- Use their own proprietary extensions to do print with CSS (not entirely inter-operable)
- Bugs/issues different in each browser/rendering engine, not easy to obtain visual feedback during document creation
- Cannot keep up with web browsers in terms of general features

Browser-as-PDF-renderer solutions

In this category:

- Pagination.JS (sophisticated, requires CSS Regions)
- SimplePagination.JS (simpler, no need for CSS Regions)

Advantages:

- Can run both headless on server and in the browser for direct visual feedback
- Lower development costs as all development is focused on JavaScript addons on top of existing rendering engines
- The CSS bugs that do exist are the same as those found in regular browsers

Issues:

- Very book specific
- Configuration of all print-related settings through JavaScript function arguments, not CSS

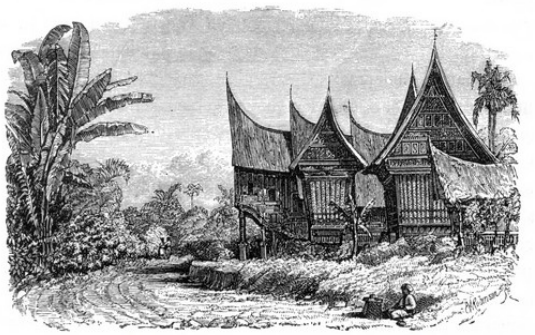
The way forward

1. Create a Browser-as-PDF-renderer solution that is
    1. generic in nature,
    2. reads configuration options by parsing CSS,
    3. Follows existing web standards
2. Participate in the creation of print-related specifications to ensure interoperability between solutions

Vivliostyle
Current status

Regular browser rendering

The Malay Archipelago



Malay Archipelago Chief's House and Rice-shed in a Sumatran Village

The Malay Archipelago is a book by the British naturalist *Alfred Russel Wallace* that chronicles his scientific exploration, during the eight-year period 1854 to 1862, of the southern portion of the *Malay Archipelago* including *Malaysia*, *Singapore*, the islands of *Indonesia*, then known as the *Dutch East Indies*, and the island of *New Guinea*. It was published in two volumes in 1869, delayed by Wallace's ill health and the work needed to describe the many specimens he brought home. The book went through ten editions in the nineteenth century; it has been reprinted many times since, and has been translated into at least eight languages.

The book described each island that he visited in turn, giving a detailed account of its *physical* and *human geography*, its volcanoes, and the variety of animals of plants that he found and collected. At the same time, he describes his experiences, the difficulties of travel, and the help he received from the different peoples that he met. The preface notes that he travelled over 14,000 miles and collected 125,660 *natural history* specimens, mostly of *insects* though also with thousands of *molluscs*, *birds*, *mammals* and *reptiles*.

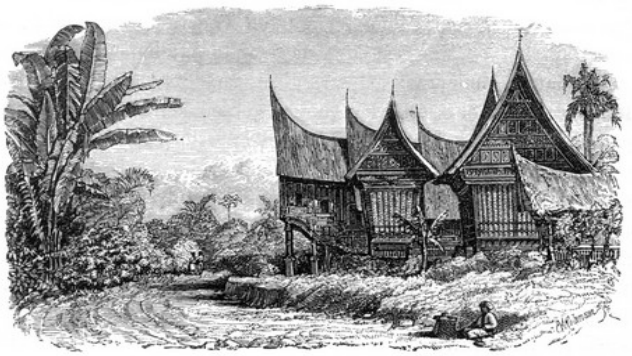
The Malay Archipelago attracted many reviews, with interest from scientific, geographic, church and general periodicals. Reviewers noted and sometimes disagreed with various of his theories, especially the division of *fauna* and *flora* along what soon became known as the *Wallace line*, *natural selection* and *uniformitarianism*. Nearly all agreed that he had provided an interesting and comprehensive account of the *geography*, natural history, and peoples of the archipelago, which was little known to their readers at the time, and that he had collected an astonishing number of specimens. The book is much cited, and is Wallace's most successful, both commercially and as a piece of literature.

Context

In 1847, Wallace and his friend *Henry Walter Bates*, both in their early twenties,Bates was 22, Wallace was 24. agreed that they would jointly make a collecting trip to the Amazon "towards solving the problem of origin of species";Mallet, Jim. "Henry Walter Bates". University College London. Retrieved December 11, 2012. *Charles Darwin*'s book on the *Origin of Species* was not published until 11 years later, in 1859, itself precipitated by a famous letter from Wallace which

Rendering using Vivliostyle.js

The Malay Archipelago



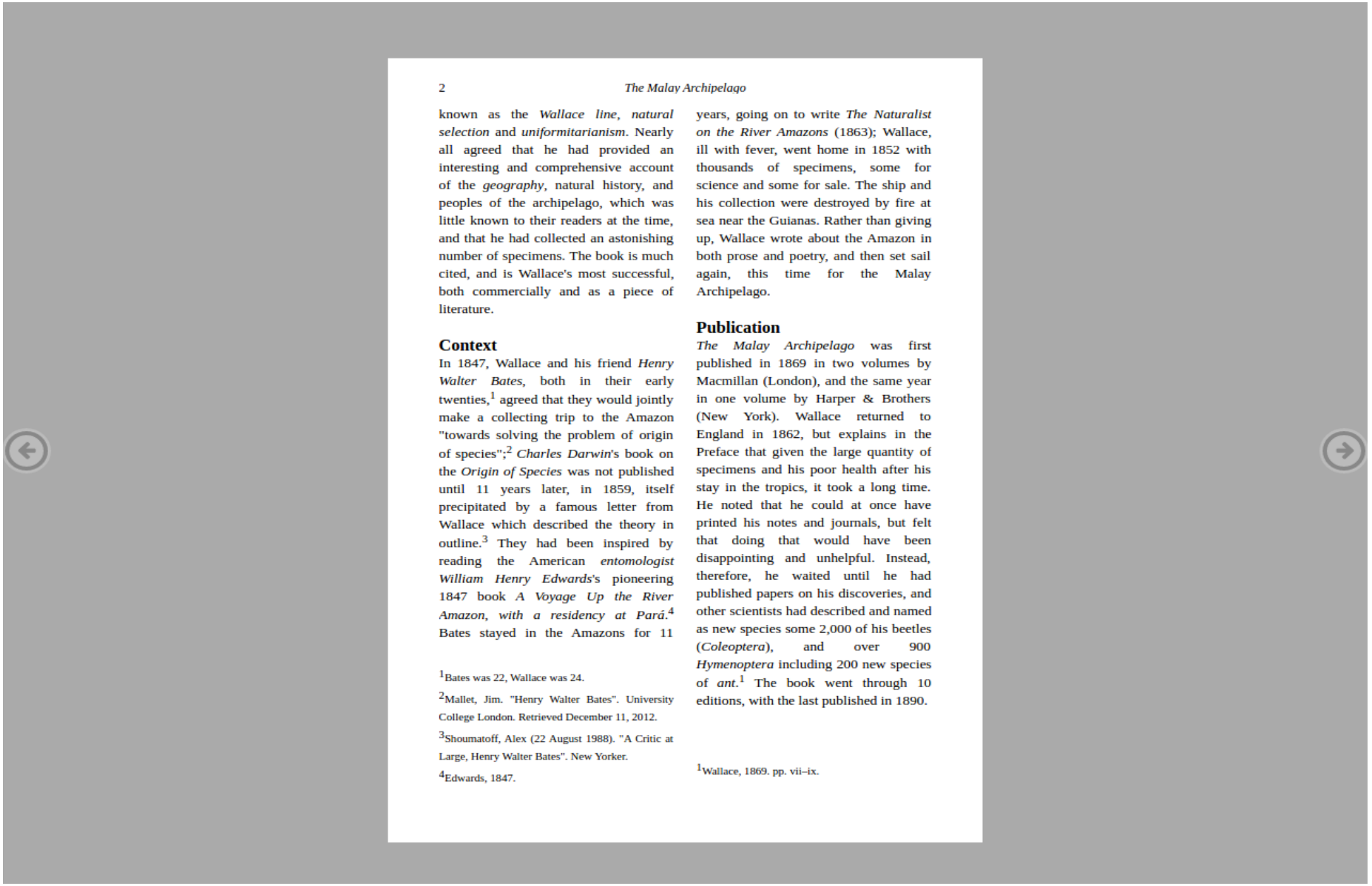
Malay Archipelago Chief's House and Rice-shed in a Sumatran Village

The Malay Archipelago is a book by the British naturalist *Alfred Russel Wallace* that chronicles his scientific exploration, during the eight-year period 1854 to 1862, of the southern portion of the *Malay Archipelago* including *Malaysia*, *Singapore*, the islands of *Indonesia*, then known as the *Dutch East Indies*, and the island of *New Guinea*. It was published in two volumes in 1869, delayed by Wallace's ill health and the work needed to describe the many specimens he brought home. The book went through ten editions in the nineteenth century; it has been reprinted many times since, and has been translated into at least eight languages.

The book described each island that he visited in turn, giving a detailed account of its *physical* and *human geography*, its volcanoes, and the variety of animals of plants that he found and collected. At the same time, he describes his experiences, the difficulties of travel, and the help he received from the different peoples that he met. The preface notes that he travelled over 14,000 miles and collected 125,660 *natural history* specimens, mostly of *insects* though also with thousands of *molluscs*, *birds*, *mammals* and *reptiles*.

The Malay Archipelago attracted many reviews, with interest from scientific, geographic, church and general periodicals. Reviewers noted and sometimes disagreed with various of his theories, especially the division of *fauna* and *flora* along what soon became

Including footnotes, headers and page numbers



known as the *Wallace line*, *natural selection* and *uniformitarianism*. Nearly all agreed that he had provided an interesting and comprehensive account of the *geography*, natural history, and peoples of the archipelago, which was little known to their readers at the time, and that he had collected an astonishing number of specimens. The book is much cited, and is Wallace's most successful, both commercially and as a piece of literature.

Context

In 1847, Wallace and his friend *Henry Walter Bates*, both in their early twenties,[1] agreed that they would jointly make a collecting trip to the Amazon "towards solving the problem of origin of species";[2] *Charles Darwin*'s book on the *Origin of Species* was not published until 11 years later, in 1859, itself precipitated by a famous letter from Wallace which described the theory in outline.[3] They had been inspired by reading the American *entomologist William Henry Edwards*'s pioneering 1847 book *A Voyage Up the River Amazon, with a residency at Pará*.[4] Bates stayed in the Amazons for 11 years, going on to write *The Naturalist on the River Amazons* (1863); Wallace, ill with fever, went home in 1852 with thousands of specimens, some for science and some for sale. The ship and his collection were destroyed by fire at sea near the Guianas. Rather than giving up, Wallace wrote about the Amazon in both prose and poetry, and then set sail again, this time for the Malay Archipelago.

[1] Bates was 22, Wallace was 24.

[2] Mallet, Jim. "Henry Walter Bates". University College London. Retrieved December 11, 2012.

[3] Shoumatoff, Alex (22 August 1988). "A Critic at Large, Henry Walter Bates". New Yorker.

[4] Edwards, 1847.

Publication

The Malay Archipelago was first published in 1869 in two volumes by Macmillan (London), and the same year in one volume by Harper & Brothers (New York). Wallace returned to England in 1862, but explains in the Preface that given the large quantity of specimens and his poor health after his stay in the tropics, it took a long time. He noted that he could at once have printed his notes and journals, but felt that doing that would have been disappointing and unhelpful. Instead, therefore, he waited until he had published papers on his discoveries, and other scientists had described and named as new species some 2,000 of his beetles (*Coleoptera*), and over 900 *Hymenoptera* including 200 new species of *ant*.[1] The book went through 10 editions, with the last published in 1890.

[1] Wallace, 1869. pp. vii–ix.

Supports non-Latin text directions

ごん狐

新美南吉

これは、私が小さいときに、村の茂平というおじいさんからきいたお話です。

むかしは、私たちの村のちかくの、中山というところに小さなお城があって、中山さまというおとのさまが、おられたそうです。

その中山から、少しはなれた山の中に、「ごん狐」という狐がいました。ごんは、一人ぼっちの小狐で、しだの一ぱいしげった森の中に穴をほって住んでいました。そして、夜でも昼でも、あたりの村へ出てきて、いたずらばかりしました。はたけへ入って芋をほりちらしたり、菜種がらの、ほしてあるのへ火をつけたり、百姓家の裏手につるしてあるとんがらしをむしりとって、いったり、いろんなことをしました。

或秋のことでした。二、三日雨がふりつづいたその間、ごんは、外へも出られなくて穴の中にしゃがんでいました。

雨があがると、ごんは、ほっとして穴からはい出ました。空はからっと晴れていて、百舌鳥の声がきんきん、ひびいていました。

ごんは、村の小川の堤まで出て来ました。あたりの、すすきの穂には、まだ雨のしずくが光っていました。川は、いつもは水が少いのですが、三日もの雨で、水が、どっとましていました。ただのときは水につかることのない、川べりのすすきや、萩の株が、黄いろくにごった水に横だおしになって、もまれています。ごんは川下の方へと、ぬかるみみちを歩いていきました。

東大寺に隣接する稲荷神社の狐像

ふと見ると、川の中に人がいて、何かやっています。ごんは、見つからないように、そうっと草の深いところへ歩きよって、そこからじっとのぞいてみました。

底本：「新美南吉童話集」岩波文庫、岩波書店　平成9年7月15日発行第2刷

Initial work on W3C specs

W3C Editor's Draft

CSS Page Floats

Editor's Draft, 29 May 2015

This version:
http://dev.w3.org/csswg/css-page-floats/
Previous Versions:
https://hg.csswg.org/drafts/raw-file/108d7e3ff204/css-page-floats/Overview.html
Feedback:
www-style@w3.org with subject line "`[css-page-floats] … message topic …`" (archives)
Issue Tracking:
Inline In Spec
Editor:
Johannes Wilm (Vivliostyle)
Former Editor:
Håkon Wium Lie (Opera Software)

Abstract

This document describes floats that move to the top or bottom of content passages. This feature has traditionally been used in print publications in which figures and photos are moved to the top or bottom of columns or pages, along with their captions. This draft describes how to achieve this effects for floats within pages, columns, regions and elements.

CSS is a language for describing the rendering of structured documents (such as HTML and XML) on screen, on paper, in speech, etc.

But will this work in the long-run?

Common argument:

> “Browser makers don't care about print, and the example of CSS Regions in Chrome shows we cannot trust them to not remove features we need”

Answer:

> “It is true that browser makers may not ever care about pages or print. But the 'Extensible Web Manifesto' [1] may help us get primitives useful to us. Even if we do not get them, browsers already contain enough general non-print features that can be reutilized for our benefit.”

[1] https://extensiblewebmanifesto.org/

What is the Extensible Web Manifesto?

- A set of principles of how the creation of web standards should change, supported by many of the members of the CSS Working Group as well as the browser creators
- Fundamental change:
  - First, stable primitives JavaScript developers should first get access to stable primitives through defined specs,
  - Later specs on a higher level are developed in close feedback cycles with JavaScript developers who can implement the new features in JavaScript as polyfills

  Print specifications
  - may end up being implemented in browsers once fully stable,
  - or they may only ever be implemented in JavaScript.

    Either way, it will allow us to do rendering for print output using common browser engines

Vivliostyle Inc.

Web: http://www.vivliostyle.com

Email: info@vivliostyle.com

Vivliostyle.js is developed by Toru Kawakubo and based on Peter Sorotokin's Adaptive Layout implementation