

Print CSS Position Paper

Creating Business Documents with HTML and CSS

Dr. Cornelia Kränzlein & Katharina Udemadu, Compart AG

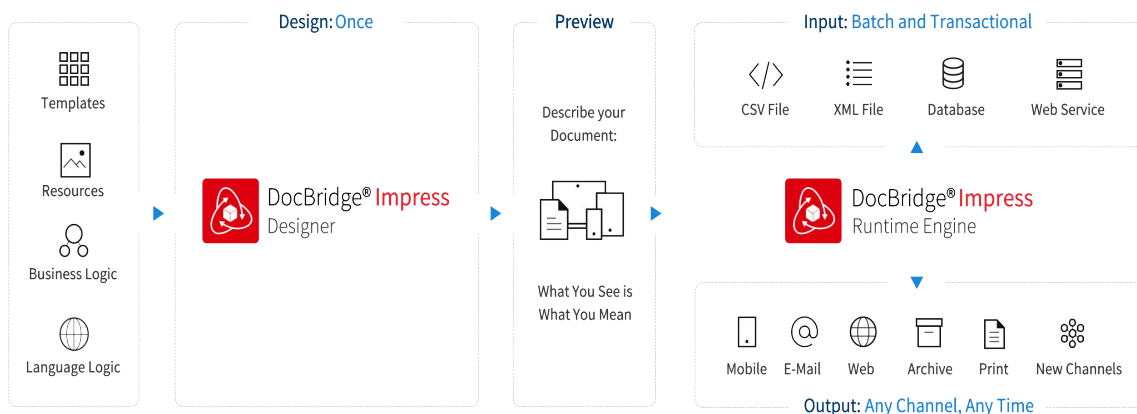
XML Prague 2020-02-13

Introduction

Compart is a leading international supplier of document and content technology solutions. The company, headquartered in Germany, has been a market presence for over twenty years and has subsidiaries in Europe and North America as well as a network of partners in Latin America.

Compart helps companies boost the efficiency of their operations with solutions for quick and flexible processing of high-volume transaction documents. The scalable and platform-independent DocBridge® family of products, designed, developed and supported by Compart, allows documents to be output anytime, anywhere, digitally or on paper.

DocBridge® Impress extends the Compart family of products by an input solution, basically allowing users to create transaction documents in HTML and to output these documents to various print and digital channels (Web, Email, Text messages).



Use Case

The creation of a transaction document such as an invoice usually consists of the following different steps. It can be compared to the setup of a mail merge with other tools.

1. Designing the document including content and layout
2. Inserting variable data
3. Setting up the output processes

Document Design

The document is created in XHTML and layouted in CSS. The document designer has to keep the different output channels in mind. Mainly two aspects must be considered:

- Are there any content differences for the different output channels, will some content only be displayed in the Web (or Email) or only output to print formats.
- Which considerations are necessary for specific output channels. Let's assume the invoice will be displayed in HTML in a customer portal and also printed as PDF for optional download and for the provider's archive system. The invoice consists of a table listing the different positions. How will the table be displayed on different devices? Will the printed table stretch across more than one page?

Variable Data

DocBridge[®] Impress inserts data from an XML data source (or an internally created XML representation of other data formats) using XSLT and XPath.

Variable data can have impact on design considerations. The invoice designer cannot know whether the invoice will consist of one or more pages. The table rows (the invoice positions) are created based on the number of positions. This means that the design for print output must be flexible. It has to include definitions for different pages types (first page of the invoice and following pages) and settings for cases where the table does not fit on one page. Do table header and footer have to be repeated? Are subtotals required?

Output Processes

Transaction documents are commonly output not once, but in different formats on different channels.

The requirements for the invoice could be:

- HTML output for Web
- PDF output for Web portal
- AFP spool for printing (one document holding all invoices)

CSS Print Requirements

Most transaction documents require one or more of the following:

Pagination: control of page breaks (inside list items and table rows), keeping heading and content together, hyphenation

Page selectors: named pages and page selectors

Headers and footers: running headers/footers, page numbers

Counters for lists or sections

Cross references to other sections or for table of contents

Meta data for different output channels and formats

Accessibility: settings required for accessible output, both Web and PDF including creation of PDF tag structure (for PDF/A, PDF/UA)

Some transaction documents also require the following:

PDF bookmarks

Leaders for table of contents or list formatting

Multi-column layout

Printer-related settings including duplex settings and tray control

Footnotes or sidenotes

Processing marks such as crop marks or folding marks

CSS Support

CSS support in the XHTML/HTML input filter and DocBridge® Impress for print documents is based on the following CSS specifications:

- Cascading Style Sheets Level 2 Revision 1 (CSS 2.1) Specification
- CSS Fonts Module Level 3
- CSS Multi-column Layout Module Level 1
- CSS Fragmentation Module Level 3
- CSS Paged Media Module Level 3
- CSS Generated Content for Paged Media Module
- CSS Generated Content Module Level 3
- CSS Text Module Level 3

Compart provides extensions for HTML and CSS for two reasons: to provide additional functionality and to cover some aspects not clearly defined in the specifications.

`xmlns:dff="http://www.compart.com/ns/dff"` namespace for proprietary elements and attributes

Use cases for dff elements include adding barcodes, inserting external documents, inserting charts, inserting query results from open data such as currency conversion, adding AFP meta data, supporting table markers.

-cp- prefix for proprietary CSS properties

Use cases for Compart CSS properties include defining/overwriting generated PDF tag structure, adding duplex information and printer tray, controlling output of table header/footer.

CSS Pain Points

Compart would like to offer customers the ability to create documents in one format for different output channels based on a specified standard.

Currently, there is a large number of CSS specifications. Many specifications also include aspects that may also be relevant for print. Information on print aspects is, however, limited.

The specification is not clear how to handle settings for repeating/not-repeating table headers/footers.

The most common print output format for HTML documents is PDF. PDF has many specific settings including tag structure, bookmarks, annotations. While there is a working draft for bookmarks, there is no information for other PDF specific settings.

The specifications primarily relevant for print settings are working drafts only, implementers are working against non-final specifications and risk implementing elements that are removed later.

What would help? CSS Print Requirements

1. Identification of a list of relevant specifications/relevant aspects of specifications for Print CSS implementers
2. Finalize relevant specifications for print (paged media module, generated content, etc.), recommendations as standard for implementers
3. Clear recommendations for accessible design, WCAG and PDF/UA
4. Review XSL-FO specifications and identify relevant features missing in Print CSS
5. Create a list of print specific issues with the CSS specifications (backlog), which implementers and other print CSS users could review and prioritize