

The print publishing experience may start with a distributor—just as someone browsing online may very well start with a search engine or aggregator of some kind. However, the retained experience is the publication I hold in my hands. The provenance information within that publication states its origin: its author and its publisher. On the Web, one leaves the search engine (by clicking a result) and is received by the publisher (who may well be the author). The reader may then return directly to that publication (as with print) and retain that same set of expectations around origin, trust, and provenance.

Web Packaging, as currently designed and proposed, seems to alter that dynamic by obfuscating the currently interrelated expression of origin, trust, and provenance.

The origin model of trust and provenance has served the Web well. However, Web Apps, Web publications, and individual Web pages have long needed the ability to be “kept” (or brought offline) for further consideration, citation, editing, archiving, and (often) eventual republication. The closest historic counterpart seems to be MHTML [RFC2557](#) built on MIME Multipart/Related [RFC2387](#). It is the foundation of any HTML email you’ve sent in the last two decades. It is also available in Chrome (behind a flag) earlier versions of Opera and Internet Explorer. It was also the foundation for the [W3C TAG’s exploration on this topic](#).

However, the lack of completed (or enabled) MHTML tooling has limited its reach. Meanwhile, formats with easier and more prevalent tooling, such as [EPUB’s Open Container Format \(OCF\)](#), are widely used, but sadly lack the ability to be identifiable as (previous) parts of the Web. Additionally, developers now frequently “freeze dry” HTML files by encoding images and fonts into `data:` URLs—hard wiring them into the HTML file itself. This has the same practical implication of “keeping a copy,” but with the similar loss of identity of both the whole and the parts.

There seem to be many use cases beyond “feeding the aggregators.” All of them pivot on bundling content while maintaining original identification. The experience of any future package format which becomes “part of the Web” will pivot, then, on how it is experienced and not simply on how it is encoded or signed or even distributed.

Benjamin Young  
Solution Architect  
John Wiley & Sons, Inc.  
Co-Chair, W3C JSON-LD WG  
Co-Editor, W3C Web Annotation

<http://bigbluehat.com/>  
<http://linkedin.com/in/benjaminyoung>