Reading Between the Lines

An Extensive Evaluation of the Security and Privacy Implications of EPUB Reading Systems

Gertjan Franken, Tom Van Goethem, Wouter Joosen

Nowadays, billions of books are digitally published based on the EPUB open technical standard. The standard's reliance on web technology translates in both a blessing and a curse, as leveraging accessible and proven web engines also implies the inheritance of their weaknesses. To shine a light on this issue, we subjected the EPUB ecosystem to a thorough evaluation.

Read our paper, published at IEEE S&P 2021

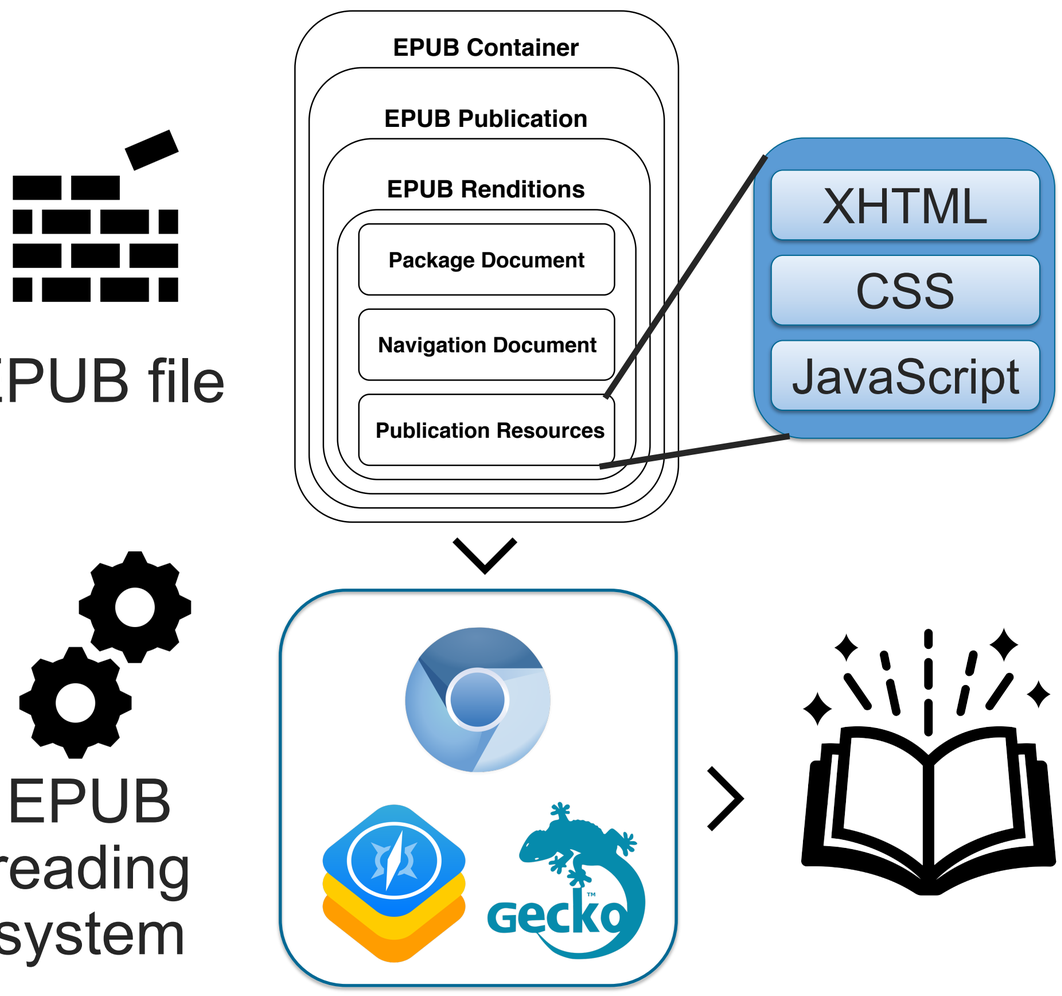
What is the state of freely available EPUB reading systems?

- Granted capabilities
- Security considerations

Are these capabilities being abused in the wild?

- Malicious EPUBs
- Tracking EPUBs

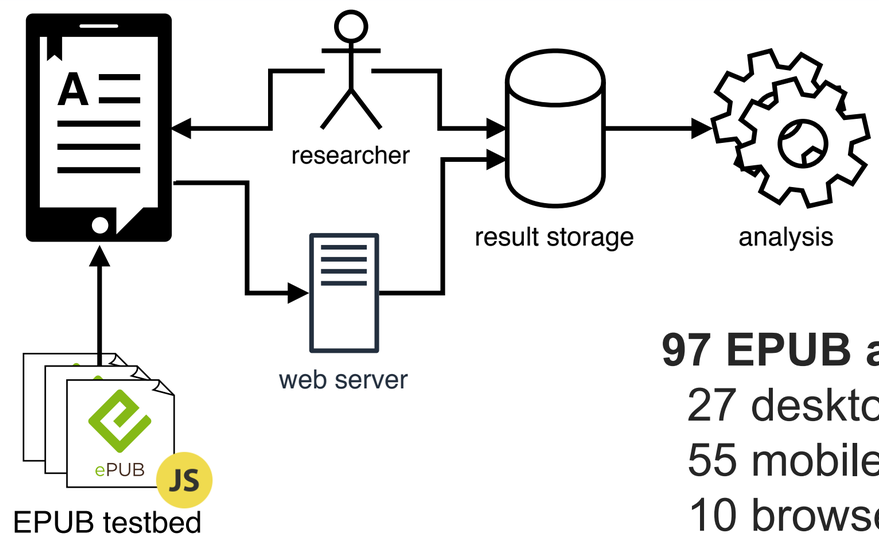
EPUB rendering



Security considerations

- User notification / consent for network activity
- Same-Origin Policy

Semi-automated EPUB evaluation testbed



97 EPUB applications
27 desktop apps
55 mobile apps
10 browser extensions
5 physical e-readers

Source code: *https://github.com/DistriNet/evil-epubs*

	Desktop	Smartphone	Browser	E-reader	Total
JavaScript execution	13 (48%)	22 (40%)	3 (30%)	1 (20%)	39 (40%)
Remote comm.	15 (56%)	20 (36%)	10 (100%)	1 (20%)	46 (47%)
Infer existence of local files	10 (37%)	6 (11%)	0	0	16 (16%)
Read content of local files	5 (19%)	3 (5%)	0	0	8 (8%)
URI handles	4 (15%)	10 (18%)	10 (100%)	0	24 (25%)
Insecure web engine	2 (7%)	0	0	1 (20%)	3 (3%)

± 50% not compliant with EPUB security considerations!

Case studies based on manual analysis

Apple Books (macOS)

Symlink validation issue
➥ Persistent DoS
➥ User information disclosure

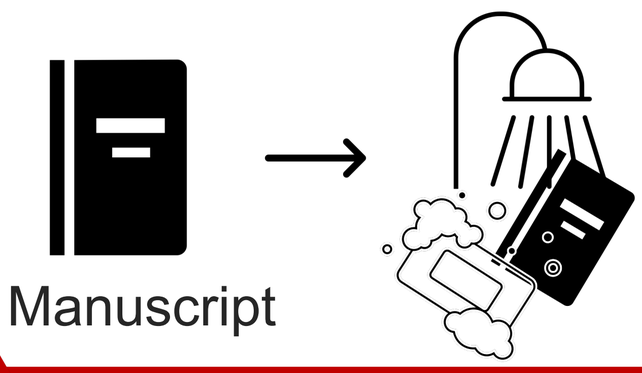
EPUBReader (Chrome, Firefox)

CSP circumvention + <all_urls> permission
➥ Universal XSS

Amazon Kindle (physical e-reader)

Input validation issue + publicly disclosed vulnerability
➥ Information leaking

Self-published EPUB sanitization



94% of EPUB self-publishing market insufficiently sanitizes!