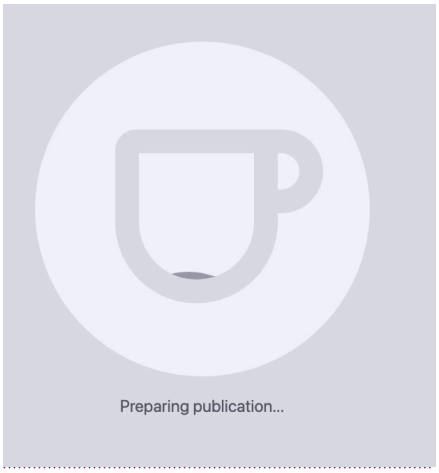
SOFTWARE

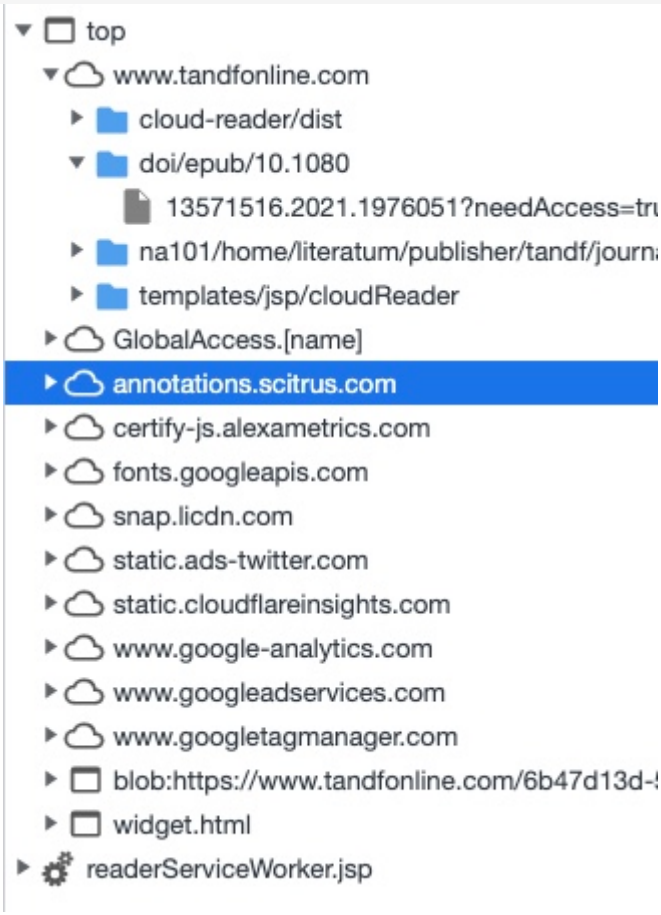
PERSONALIZED PDFS

22.01.2022

Scientific publishers are creating now more and more dynamic PDFs. Why do we know? There is an unexpected loading delay of a PDF from Routledge / Taylor & Francis group that I observed recently. First I thought about some DDos protection, but is indeed a personalized document.



These websites are all being contacted while creating this PDF:



Scitrus.com seems to be part of a larger reference organizer network and links to science-connect.io. Alexametric.com is the soon to be retired Alexa internet / Amazon service. Snap.lidcdn.com forwards to px.ads.linkedin.com, the business social network. Then we have Twitter ads, Cloudflare security and Google Analytics. All major players now know that my IP is interested in COVID-19 research. Did I ever agree to submit my IP and time stamp when looking up a rather crude scientific paper?

This is exactly what the German DFG already warned us about last October

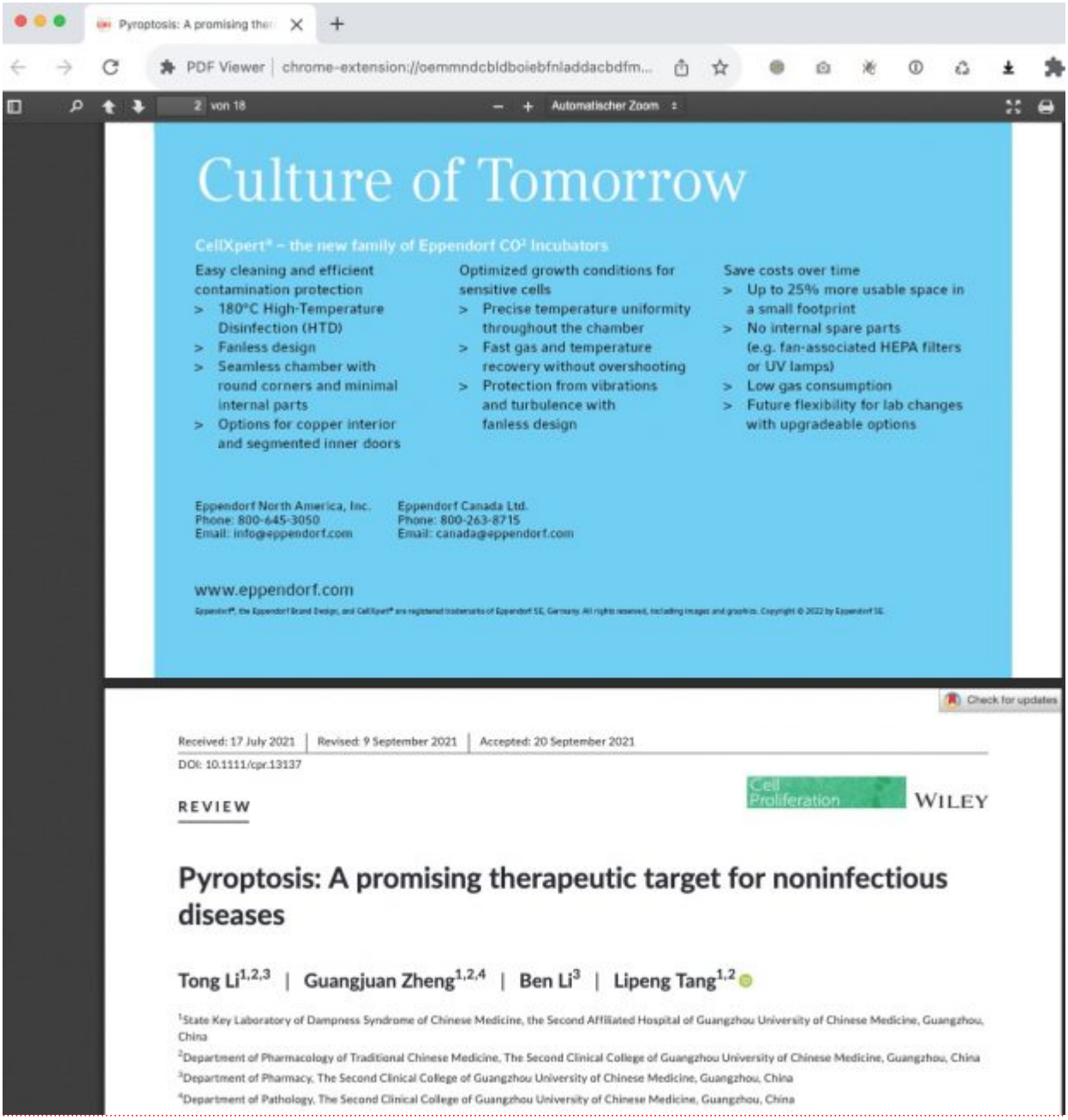
> For some time now, the major academic publishers have been fundamentally changing their business model with significant implications for research: aggregation and the reuse or resale of user traces have become relevant aspects of their business. Some publishers now explicitly regard themselves as information analysis specialists. Their business model is shifting from content provision to data analytics.

Another paper describes the situation as “Forced marriages and bastards”…

My question is : Will Francis & Taylor even do more? The structure of PDFs allows including objects including Javascript. When examining “document.pdf” using pdf-parser I could

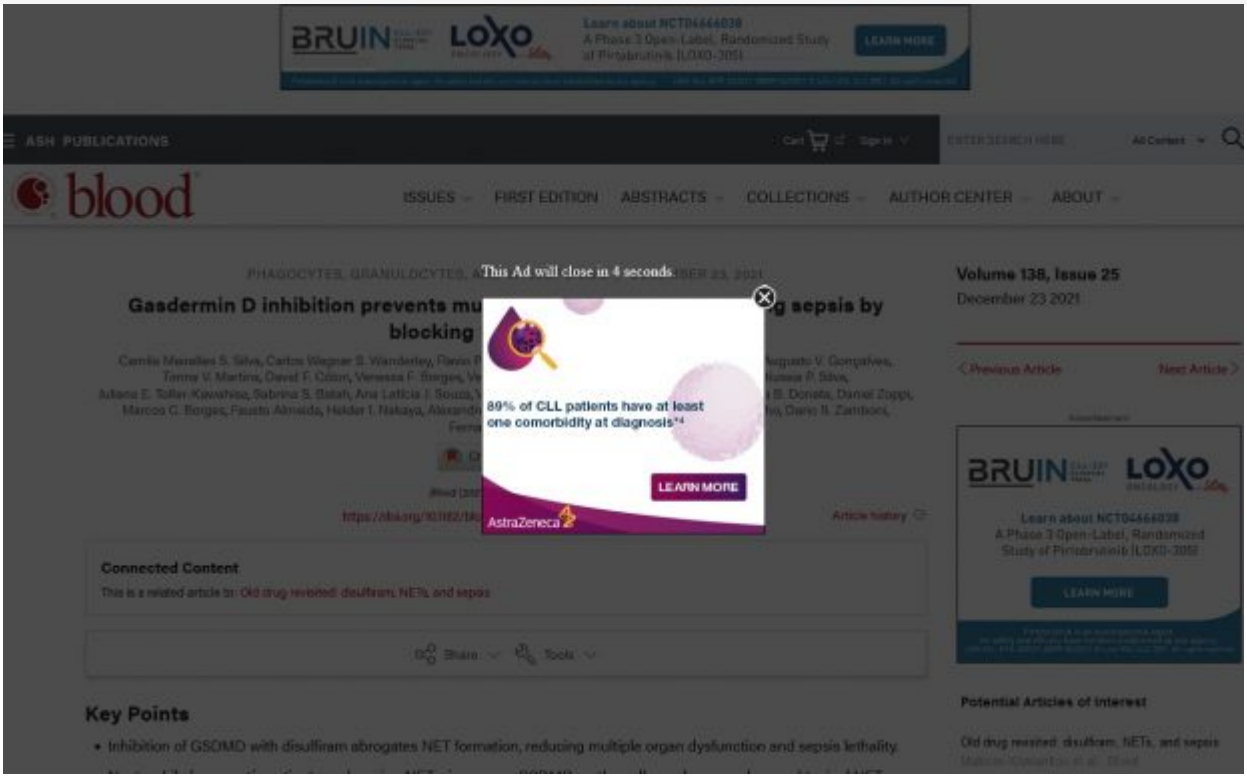
not find any javascript or my current IP in clear text. I cannot exclude however that the chopped up IP is stamped somewhere in the document. So I will have try again at a later time point and redo a bitwise analysis. of the same PDF delivered on another day.

At least the DFG document says that organisations might argue that such software allows for the prosecution of users of shadow libraries. While I have doubts that this is legal, we already see targeted advertisement as I received this PDF from Wiley that included an Eppendorf ad.



Screenshot 20.1.2022

When I downloaded this document a second time using a different IP it was however identical. Blood/Elsevier only let's you even download only after watching a small slideshow…



Screenshot 20.1.2022