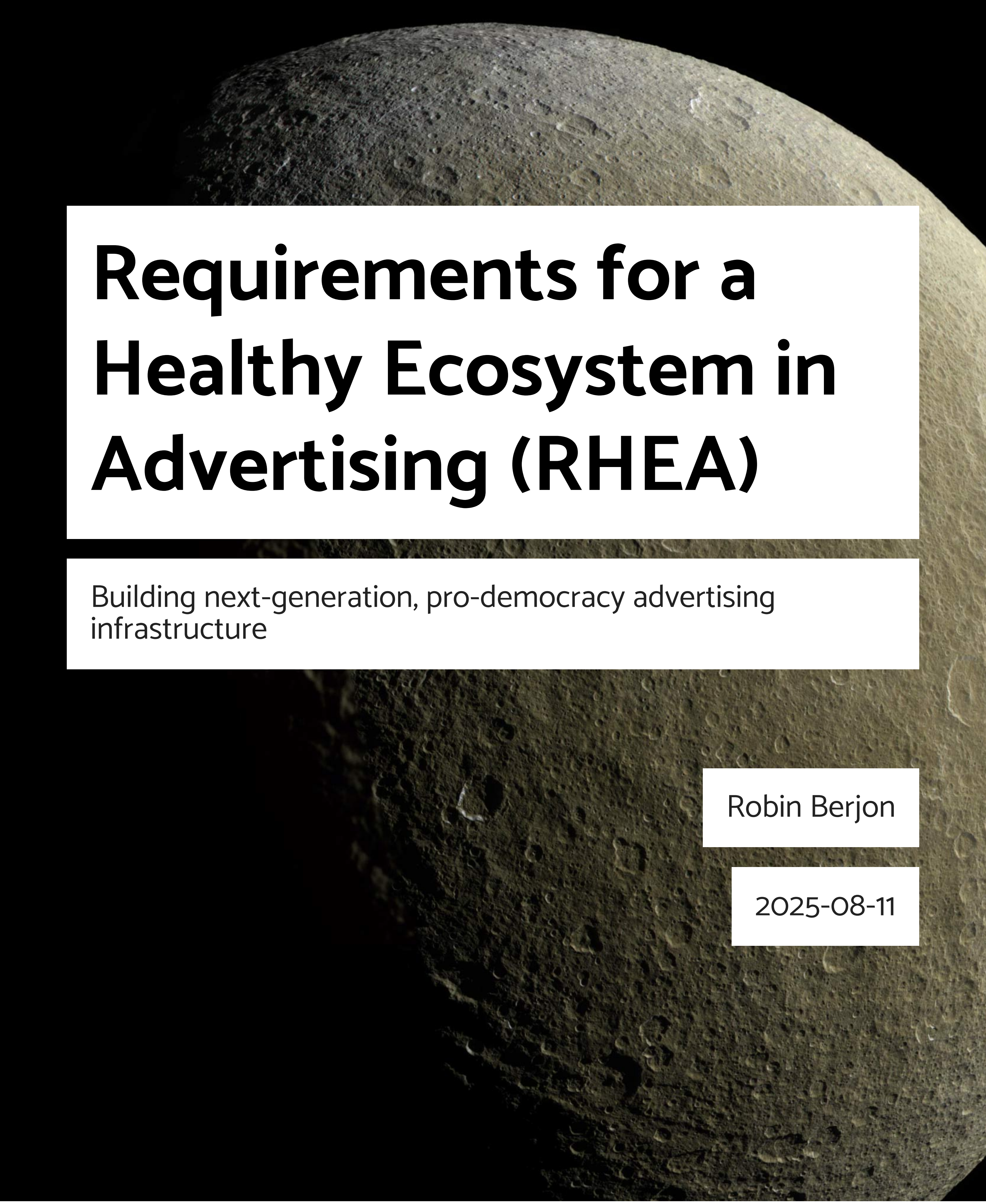

Requirements for a Healthy Ecosystem in Advertising (RHEA)

Building next-generation, pro-democracy advertising infrastructure

Robin Berjon

2025-08-11

supramundane *agency*.

Advertising is a critical input to much of the digital economy, not least of which media. In fact, it has proven essential to the point of being described as "the business model of the internet." Unfortunately, however, advertising has not been governed with the care that such an important piece of infrastructure deserves. Advertising is unpopular with technologists who want to do good, activists and civil society organisations often simply propose to eliminate it or limit themselves to superficial proposals that don't engage with the technology or economics (e.g. "just use contextual"), and regulators have found the advertising system highly opaque and complex to engage with. It is common that publishers and marketers would only partially understand the system they use, which in turn stokes fear of what may happen if anything were to change. The field is highly concentrated around a Google/Meta duopoly and fraught with a long list of ills, but a comprehensive vision for a stakeholder-centric alternative has yet to emerge.

This document outlines what an alternative could be. It doesn't claim to be complete and certainly not detailed, but it offers a path forward that can hopefully be iterated upon and deployed.

An Ecosystem Of Multifaceted Problems

There is extensive literature about issues in digital advertising, and summarising it is beyond the scope of this paper. Instead, we offer a list of the more salient problems in the space so as to inform both the urgent need for reform and the shape of the solution space we consider further down.

Taxing The Media to Subsidise Disinformation

Legacy digital advertising is a system that, at the scale of the entire internet, taxes high-quality contexts (often the media) and redistributes that money to slop farms and disinformation operators, while advertising intermediaries take a cut.

To understand how that works, it is important to understand that today's digital advertising shares data in such a way that the same person will be recognised by intermediaries across completely unrelated contexts. Crucially, this means that a person can be observed in one context and targeted in another.

Not all contexts are equal. Quality publishers (of all sizes) develop specific audiences over time and through hard work. These audiences are valuable for advertisers because they correspond to a meaningful group of people that can correlate well with brand goals or a given purchase intent. Slop farms and disinformation networks, on the other hand, do not produce valuable, well-developed audiences. Instead, they try to drive traffic up through any means available. In addition to the low quality of the context they offer, they also have amorphous audiences without clear advertising value. When contexts are isolated (and also when advertisers have better control over where their advertising ends up), it is more challenging for slop and disinformation to be profitable.

This all changes when people can be targeted across contexts. A given person can be observed as being part of a high-value audience and later targeted in a low-value context. This both decreases the valuation of the quality context — since it is getting fewer, lower bids for its own audience — and increases the revenue that flows to slop and disinformation, what is known as *Made For Advertising (MFA)* sites. It is a distributional arrangement that takes money from hard-working media to subsidise disinformation. This is detrimental to the media, to democracy, and to the information ecosystem overall.

A few additional details add some colour. First, slop is profitable on volume (of which there is a lot) and also provides an outlet for lower-quality ads. There is an equilibrium point that blends a given proportion of quality and slop contexts at which an advertising marketplace such as Google's is maximally profitable. Since the marketplace controls data distribution and prices, they are in a position to ensure this equilibrium is maintained.

Second, it's tempting to believe that marketers may benefit from lower prices thanks to this but it's not obvious that they do — they often have poor visibility into the contexts in which their advertising reaches their intended clients.

And finally, it's important to keep in mind that what policy documents from Google or Meta describe as "publishers" is a large undifferentiated group that almost always includes active disinformation networks and slop farms.

A pro-democracy solution must put an end to this subsidy system, both to make quality publishers more profitable and disinformation a less attractive business.

Privacy

Issues of surveillance in digital advertising are well established. The loading of one single ad typically involves broadcasting someone's interaction with a given page or system to hundreds of additional parties, and after the ad has been loaded yet more parties are contacted. All of these actors obtain identifiers that enable them to track people's behaviour across their entire online presence. Both publishers and marketers are typically unable to provide a reliable list of who has seen any given person's data, and are even less able to account for processing purposes. Regulatory interventions have produced very little impact and consent-based approaches often inconvenience people without affecting data collectors much.

There is at best only limited evidence to support the idea that this data collection and processing is needed. Not one single reproducible study has been produced, results point in different directions depending on method and data set, and many studies are conflicted (or don't declare conflicts), fail to test meaningful counterfactuals, and some commonly-cited ones are preposterously wrong.

There is, however, an advantage to surveilling people in many contexts: it's an effective way to predict that a given person is about to purchase a specific product (or carry out a similar action), show them an ad right before they buy it, and claim the conversion. This is worth unpacking. Marketers are keen to measure the effectiveness of their campaigns by attributing conversions — a specific action taken by a person, often a sale but not necessarily — to whoever showed that person an ad that "triggered" this conversion. But deciding what ad might *cause* which given person at what given time in which specific context to buy something is inordinately hard. It is *a lot* easier to detect that a given person is about to make a purchase no matter what — especially if you have data about what they search for through a search engine, their entire browsing history through your browser, their location through your mapping application and a mobile OS, and access to what they are emailing about.

Showing an ad for a product that someone is about to buy anyway is a net loss for the advertiser: that is literally the *last* person who should be seeing that ad. But the advertiser doesn't know that, and it makes it possible to claim the conversion, which means greater advertising dominance and more future spend. There is evidence that this practice is taking place ([*hbr-targeting*]) and no study has eliminated the concern. The massive data leak that is digital advertising is therefore not known to be better for advertisers and, as deployed, might be worse for them. We already know from the previous section that it does not help publishers.

Fraud

In addition to intermediary practices that are detrimental to people, publishers, and marketers, the digital advertising environment is riddled with fraud. A report from the World Federation of Advertisers speaks for itself ([*wfa-ad-fraud*]). They predict ad fraud to be worth USD $50bn in 2025 at the conservative end of the spectrum (and up to almost three times that), potentially making it second only to the cocaine and opiates market as a form of organised crime.

Interestingly, the WFA estimates that whenever ad fraud is committed, less than half of the money goes to the fraudster, with the rest flowing to "legitimate" intermediary actors. One doesn't need to be cynical to note that the revenue amounts and the widespread nature of fraud do not incentivise intermediaries to crack down on it.

To make matters more interesting, the suggested cure for these problems can often make the problem worse, for instance when Invalid Traffic (IVT) or brand safety vendors (who are supposed to keep ads from appearing in wrong, including fraudulent, contexts) systematically defund climate reporting or use their tooling to index publishers' content and compete against them for contextual targeting.

Security

The technical practices of adtech and martech are also intrinsically insecure. The entire edifice is built on top of unverified third-party code injection. When (not if) a marketing vendor is compromised, the attacker becomes able to run code on often tens of thousands of websites as if they were the website operator, allowing them to interact with backend APIs or to capture data entered by users, up to and including username/password combinations.

This is a direct consequence of the open-ended promiscuity of the adtech stack, as empowered by browser vendors. It is, in a very concrete sense, strictly impossible for a publisher to abide by their legal obligations to be diligent with cybersecurity if they also wish to display advertising from today's industry.

This additionally makes data breaches impossible to audit after the fact. In many cases, there is simply no way to know which third parties had code injected into a page as a breach was ongoing.

Whenever a security bug is uncovered in a browser, advertising provides an excellent delivery mechanism for it. And because code is injected dynamically, it is possible for an attacker to inject malicious code for a time, attain a given objective, and then remove the malicious code without leaving any trace.

Espionage

It is tempting to lump all issues downstream of data collection under privacy, but when you combine the scale of the surveillance, the sensitive nature of the data including all online activity and geolocation, and a technical architecture that entirely obfuscates who is listening you obtain a system with societal and geopolitical concerns that exceed privacy issues typically more centred on furthering individual agency.

Indeed, it is possible to buy information off the shelf about military and intelligence workers who attend brothels and visit nuclear vaults ([*wired-nuclear-brothels*]) and thanks to the wide sharing of identifiers between data brokers the U.S. government can create a one-stop-shop for sensitive personal data provided by the private sector ([*spy-one-stop-shop*]). This confirms earlier work from The New York Times in which reporters, with access to only a *tiny* subset of the data that brokers have were able to track people to nuclear plants, church, family planning ([*nyt-last-night*]).

Nothing about this data limits its use to the espionage of nation-states, it is perfectly applicable to corporate espionage as well. Listing all the people who work in a given building during the day and leave digital trails that expose them to blackmail is trivial. As is knowing what employees are reading about on the internet, what they are buying, or where they are travelling to.

While not specific to advertising, this issue is almost entirely driven by advertising-related data collection.

Opacity

The digital advertising environment is so opaque that very little evidence-based decision-making is possible. Marketers track campaign performance as best they can and make the most informed decisions that they are able to (e.g. using mixed-media modelling) but by and large they lack the ability to sufficiently understand which levers are really effective and where their campaigns are really shown.

The biggest impact from this is risk aversion. Actors will have an ad strategy that delivers *something* and any major change feels like it could bring everything down. It is so difficult to meaningfully attribute causality — even if experienced practitioners know how to make localised improvements — that it is impossible to assess changes without tests and very difficult to create meaningful tests.

Advertisers are often in the dark about where their ads were placed. Publishers are often in the dark about which impression drove which income, and completely in the dark about most ads they carry or how much the buyer paid for them. Calculating the adtech tax is at best challenging and some of the better attempts at tracing money from buyers to sellers failed to account for very significant sums.

Concentration and Governance

There is little need to expand on problems of concentration in the advertising market now that Google has been convicted as a monopolist in *both* digital advertising in general and search advertising. The overall digital advertising market has a large number of small actors because there are plenty of services — many of them of dubious value — that can be overlaid atop the core money-makers, but there is no doubt as to who dominates the market and makes the rules. Even Meta, the other arm of the duopoly, takes second place to Google's global empire.

While this situation is well documented, it is worth underlining a couple of specificities that should guide efforts to fix the advertising market.

First, as explained at length by Dina Srinivasan, "*Google dominates advertising markets by engaging in conduct that lawmakers prohibit in other electronic trading markets: Google's exchange shares superior trading information and speed with the Google-owned intermediaries, Google steers buy and sell orders to its exchange and websites (Search & YouTube), and Google abuses its access to inside information.*" ([why-google-dominates-ads])

That is possible because, unlike other electronic markets (and, in fact, markets in general) advertising trading markets are largely unregulated or operate under unenforced regulation. This is possible in part because the digital advertising industry was granted the leeway to self-regulate for 25 years (which it largely did not do), and in part because Google controls and sets the rules for the marketplaces, data ingress points (browsers and operating system), and largest market participants. Google is the actor with the most power and the most to lose from sharing governance over advertising markets.

A second consideration is that it isn't possible to completely separate the already-complex digital advertising system from related considerations involving browsers, operating systems, and search.

Part of the issue has to do with privacy. For one, Google Chrome is not only the only remaining significant browser to offer no privacy protection but it also deploys an impressive array of deceptive patterns to trick its users into disclosing their entire browsing history to Google — something which Google then uses to optimise its search engine, an advertising business. Google isn't the only one to blame: Apple also invented mobile ad IDs so that people would be tracked in apps and markets privacy but defaults to Google Search.

But, more importantly, the manner in which Google's search business maintains its dominance has a direct impact on the value of advertising across the digital sphere, defunding other actors. In a nutshell, money that is spent on search advertising isn't spent elsewhere in the advertising system. But, as established in the U.S. Google Search case, Google's monopoly in search allows it to charge supracompetitive prices in search, and the centrality of search means buyers cannot easily switch to spending elsewhere. This means that Google Search is depreciating ad prices for the rest of the digital sphere. In turn, that search monopoly is enforced by browser vendors and mobile operating systems, who get a cut of that revenue in exchange for the service. The system operates as a de facto cartel, making money by artificially driving up the value of Google Search ads and defunding ad revenue for the rest of the web.

Sustainability

The real-time nature of ad bidding requires having every ad request processed by hundreds of parties. This is highly inefficient. The vast amounts of data being collected, stored, access, and processed for this system also incur energetic costs.

Additionally, ad creatives are often exceedingly heavy (in bytes), which uses up bandwidth and processing power. Since they are typically seen thousands if not millions of times, the effect adds up.

What's more, some types of ad fraud can be particularly processing-intensive.

Overall, while sustainability considerations in advertising have rarely been foregrounded, there is meaningful room for improvement along this dimension as well.

Our Goals: Who To Improve Advertising For?

tk

- □ Better for publishers: describe the subsidy system and how we need to end it
    - □ for each go through all the intro problems
- □ Better for advertisers: traceability problems, lots of marketing options are bad
- □ Better for people: credible privacy
- □ Governed ad safety
- □ from RHEA
- □ Exemptions + technical guarantees for the processing
- □ Explain that this is intended to help publishers too
- □ Single controller operations (apart perhaps from some very limited technical services like serving)
- □ Article 40/41
- □ Could consider using the Data Act and data exchanges (?) to support some form of targeting
    - □ Use sortition-based privacy (see Asia, also Max von G.)
- □ Beckn specialised for this
- □ SDAs — why this makes them *actually* possible
- □ Some geo
- □ PPA
- □ Fraud prevention
- □ KYC requirements
- □ Money: can this use a digital currency and protocol that supports splitting well?
- □

Locked down, but that's okay because purchases are made in small batches

- □ Content-addressed and all content is available at purchase time (this allows extra charges for creative size to be addressed)
- □ Bring back some GARUDA
- □ I don't think that we want browsers involved
- □ Split the components of the architecture so that we can have different governance for different elements — AT-style

Privacy

tk

Marketplace

tk

Creatives

tk

Governance

tk

References

[hbr-targeting]

Does Personalized Advertising Work as Well as Tech Companies Claim?. Bart de Langhe; Stefano Puntoni. 2021-12-16. https://hbr.org/2021/12/does-personalized-advertising-work-as-well-as-tech-companies-claim

[nyt-last-night]

Your Apps Know Where You Were Last Night, and They’re Not Keeping It Secret. Jennifer Valentino De Vries, Natasha Singer, Michael H. Keller And Aaron Krolik. 2018-12-10. https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html

[spy-one-stop-shop]

U.S. Spy Agencies Are Getting a One-Stop Shop to Buy Your Most Sensitive Personal Data. Sam Biddle. 2025-05-22. https://theintercept.com/2025/05/22/intel-agencies-buying-data-portal-privacy/

[wfa-ad-fraud]

Compendium of ad fraud knowledge for media investors. World Federation of Advertisers; The Advertising Fraud Council. 2016. https://swa-asa.ch/wAssets/docs/publikationen/de/branchenempfehlungen-swa/WFACompendiumofAdFraudKnowledge.pdf

[why-google-dominates-ads]

Why Google Dominates Advertising Markets. Dina Srinivasan. 2019-12-09. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3500919

[wired-nuclear-brothels]

Anyone Can Buy Data Tracking US Soldiers and Spies to Nuclear Vaults and Brothels in Germany. Dhruv Mehrotra; Dell Cameron. 2024-11-19. https://www.wired.com/story/phone-data-us-soldiers-spies-nuclear-germany/