Report from Dr Johnny Ryan – Behavioural advertising and personal data

Background and expertise

My name is Johnny Ryan. I am the Chief Policy and Industry Relations Officer for Brave, a privacy-focussed Internet Browser.

I have worked on both sides of the ad tech and publisher divide. Before I joined Brave I was responsible for research and analysis at PageFair, an advertising technology company. In that role, I participated in standards setting working groups for the ad tech industry. In a previous role, before PageFair, I worked at The Irish Times, a newspaper, where I was the Chief Innovation Officer.

I have had other roles, in academia and in policy. I am the author of two books on Internet issues. One is a history of the technology, which has featured on the reading list at Harvard and Stanford. The other was the most cited source in the European Commission's impact assessment that decided against pursuing Web censorship across the European Union. I am a Fellow of the Royal Historical Society, and a member of the World Economic Forum's expert network on media, entertainment and information.

I have a PhD from the University of Cambridge, where I studied the spread of militant memes on the Web.

My expert commentary on the online media and advertising industry has appeared in The New York Times, The Economist, The Financial Times, Wired, Le Monde, NPR, Advertising Age, Fortune, Business Week, the BBC, Sky News, and various others.

How personal data are used in behavioural online advertising.

Every time a "behaviourally" targeted advert is served to a person visiting a website, the system that selects what advert[1] to show that person broadcasts their personal data to hundreds or thousands of companies.

These personal data include the URL of every page a user is visiting, their IP address (from which geographical position may be inferred), details of their device, and various unique IDs that may have been stored about the user previously to help build up a long term profile about him or her.

[1] This system is known as "Real-time bidding", or sometimes referred to as "programmatic" (which simply means automatic) advertising.

It is also interesting to note that this system is a relatively recent development in online media. Only as recently as December 2010 did a consortium[2] of advertising technology ("AdTech") companies agree the methodology for this approach to tracking and advertising. Before this, online advertising was placed by far more simple ad networks that sold ad slots on websites, or by highly lucrative direct sales deals by publishers.[3]

As detailed below, despite the grace period leading up to the GDPR, the AdTech industry has built no adequate controls to enforce data protection among the many companies that receive data.

How personal data are "broadcast".

A large part of the online media and advertising industry uses a system called "RTB", which stands for "real time bidding". There are two versions of RTB.

- "OpenRTB" is used by most significant companies in the online media and advertising industry.
- "Authorized Buyers", Google's proprietary RTB system. It was recently rebranded from "DoubleClick Ad Exchange" (known as "AdX") to "Authorized Buyers".[4]

Note that Google uses both OpenRTB and its own proprietary "Authorized Buyers" system.[5]

---

[2] The consortium included DataXu, MediaMath, Turn, Admeld, PubMatic, and The Rubicon Project. See a note on the history of OpenRTB in "OpenRTB API Specification Version 2.4, final draft", IAB Tech Lab, March 2016 (URL: https://www.iab.com/wp-content/uploads/2016/03/OpenRTB-API-Specification-Version-2-4-FINAL.pdf), p. 2-3.

[3] Only in 2006 did the first "ad exchange" emerge, and enable ad networks to auction space on their clients' websites to prospective buyers. A pioneer was Right Media, which was bought by Yahoo!. "RMX Direct: alternative ad networks battle for your blog", Tech Crunch, 12 August 2006 (URL: https://techcrunch.com/2006/08/12/rmx-direct-alternative-ad-networks-battle-for-your-blog/?_ga=2.239524803.1716001118.1536329047-1016164068.1536329047)

[4] "Introducing Authorized Buyers", Authorized Buyers, Google (URL: https://support.google.com/adxbuyer/answer/9070822, retrieved 24 August 2018).

[5] "OpenRTB Integration", Authorized Buyers, Google (URL: https://developers.google.com/authorized-buyers/rtb/openrtb-guide, retrieved 24 August 2018).

The OpenRTB specification documents are publicly available from the New York-based IAB TechLab.[6] The "Authorized Buyers" specification documents are publicly available from Google.

Both sets of documents reveal that every time a person loads a page on a website that uses real-time bidding advertising, personal data about them are broadcast to tens - or hundreds - of companies. Here is a sample of the personal data broadcast.

- What you are reading or watching
- Your location (OpenRTB also includes full IP address)
- Description of your device
- Unique tracking IDs or a "cookie match" to allow advertising technology companies to try to identify you the next time you are seen, so that a long-term profile can be built or consolidated with offline data about you
- Your IP address (depending on the version of "RTB" system)
- Data broker segment ID, if available. This could denote things like your income bracket, age and gender, habits, social media influence, ethnicity, sexual orientation, religion, political leaning, etc. (depending on the version of "RTB" system)

These data show what the person is watching and reading, and can include - or be matched with - data brokers' segment IDs that categorise what kind of people they are.

A more complete summary of the personal data in Open RTB bid requests, which are used by all RTB advertising companies, including Google, is provided for your convenience in Appendix 1.

A summary of the personal data in Google's proprietary bid requests is provided in Appendix 2.

Relevant excerpts from the OpenRTB "AdCOM" specification documents are presented in Appendix 3, and excerpts from Google's proprietary RTB specification documents are provided in Appendix 4.

How it works

A diagram of the flow of information is provided below.

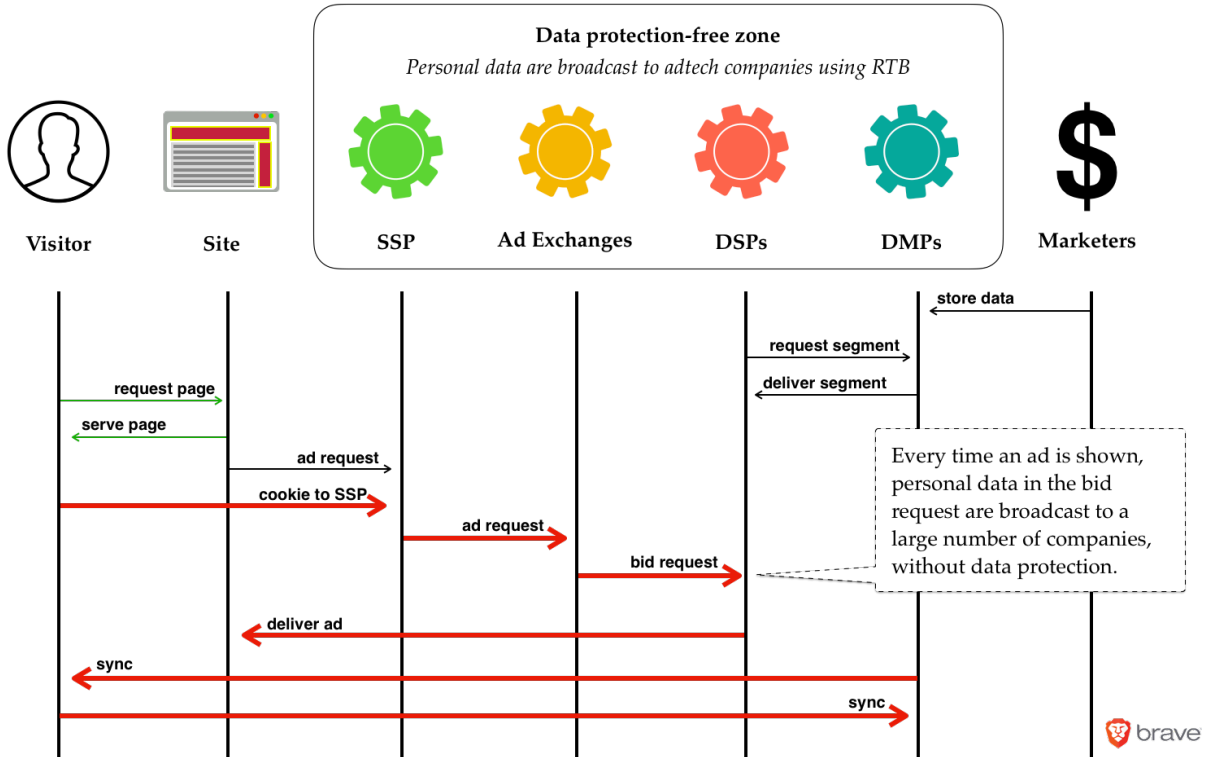
In summary, the broadcast of these personal data under RTB is referred to as an "RTB bid request". This is generally broadcast widely, since the objective is to solicit bids from companies that might want to show an ad to the person who has just

[6] The IAB is the standards body and trade lobby group of the global advertising technology industry. All significant ad tech companies are members. The IAB has local franchises across the globe. Its standards-setting organisation is IAB TechLab.

loaded the webpage. An RTB bid request is broadcast on behalf of websites by companies known as "supply side platforms" (SSPs) and by "ad exchanges".

The diagram below shows how personal data are broadcast in bid requests to multiple Demand Side Partners (DSPs), which then decide whether to place bids for the opportunity to show an ad to the person in question. The DSP acts on behalf of an advertiser, and decides when to bid based on the profile of person that the advertiser has instructed it to target.

Sometimes, Data Management Platforms (DMPs), of which Cambridge Analytica is a notorious example, can perform a "sync" that uses this personal data to contribute to their existing profiles of the person. In it worth noting that this sync would not be possible without the initial bid request.



The overriding commercial incentive for many ad tech companies is to share as much data with as many partners as possible, and to share it with partner or parent companies that run data brokerages. Clearly, releasing personal data into such an environment has high risk.

Despite this high risk, RTB establishes no control over what happens to these personal data once an SSP or ad exchange broadcasts a "bid request". Even if bid request traffic is secure, there are no technical measures that prevent the recipient of a bid request from, for example, combining them with other data to create a profile, or from selling the data on. In other words, there is no data protection.

That IAB Europe's own documentation for its "GDPR Transparency & Consent Framework", says that a company that receives personal data should only share these data with other companies if it has "a justified basis for relying on that Vendor's having a legal basis for processing the personal data".[7] In other words, the industry is adopting a "trust everyone" approach to the protection of very intimate data once they are broadcast.

There are no technical measures in place to adequately protect the data. I note that IAB Europe recently announced that it is developing a tool, in collaboration with an organisation called The Media Trust, that will attempt to determine whether the "consent management platforms" (CMPs) that participate in the IAB Europe Framework are complying with the Framework's policies. According to IAB Europe's press release, the tool "validates whether a CMP's code conforms to the technical specifications and protocols detailed in the IAB Europe Transparency & Consent Framework".[8]

But the tool, which is currently only in beta, will be inadequate to protect personal intimate personal data broadcast in bid requests. This is because - even if it could police all web-based data transmission[9] - it would still have no way of knowing whether, for example, a company had set up a continuous server to server transfer of personal data to other companies.

Once the personal data are released in a bid request to a large number of companies, the game is over. In other words, once DSPs receive personal data they can freely trade these personal data with business partners, however they wish.

This is particularly egregious since the data concerned are very likely to be "special categories" of personal data. The personal data in question reveal what a person is watching online, and often reveal specific location. These alone would reveal a person's sexual orientation, religious belief, political leaning, or ethnicity. In addition, a "segment ID" that denotes what category of person a data broker or other long-term profiler has discovered a person fits in to.

---

[7] "IAB Europe Transparency & Consent Framework – Policies", IAB Europe, 25 April 2018 (URL: http://www.iabeurope.eu/tcfdocuments/documents/legal/currenttcfpolicyFINAL.pdf), p. 7.

[8] "IAB Europe Press Release: IAB Europe CMP Validator Helps CMPs Align with Transparency & Consent Framework", IAB Europe, 12 September 2018 (URL: https://www.iabeurope.eu/all-news/press-releases/iab-europe-press-release-iab-europe-cmp-validator-helps-cmps-align-with-transparency-consent-framework/).

[9] See "Data compliance", The Media Trust website (URL: https://mediatrust.com/how-we-help/data-compliance)

Moreover, the industry concerned is aware of the shortcomings of this approach, and has continued to pursue it regardless.

RTB bid requests do not necessarily need to contain personal data. If all industry actors agreed, and amended the standard under the stewardship of the IAB, then bid requests that contain no personal data could be passed between ad tech companies to target relevant advertising by general context. This, however, would prevent these companies and their business partners from building profiles of people, which would have a revenue implication. The industry is currently finalising a new RTB specification (OpenRTB 3.0), which continues to broadcast personal data without protection in the same way that previous versions of the OpenRTB system. Tables from OpenRTB 3.0 that show the personal data in question are presented for your convenience in Appendix 4.

Online advertising that uses this approach will continue to disseminate details about what every person is reading or watching in a constant broadcast to a large number of companies. These personal data are not protected. This dissemination is continuous, happening on virtually every website, every single time a person loads a page.

This is a widespread and troubling practice. The scope of the industry affects the fundamental rights of virtually every person that uses the Internet in Europe.

Concerns about these practices (news reports, NGO investigations, regulatory consideration etc.)

Survey data over several years demonstrates a general and widespread concern about these practices. The UK Information Commissioner's Office's own survey, published in August 2018, reports that 53% of British adults are concerned about "online activity being tracked".[10]

In 2017, GFK was commissioned by IAB Europe (the AdTech industry's own trade body) to survey 11,000 people across the EU about their attitudes to online media and advertising. GFK reported that only "20% would be happy for their data to be shared with third parties for advertising purposes".[11] This tallies closely with survey that GFK conducted in the United States in 2014, which found that "7 out of 10 Baby

[10] "Information rights strategic plan: trust and confidence", Harris Interactive for the Information Commissioner's Office, August 2018, p. 21.

[11] "Europe online: an experience driven by advertising. Summary results", IAB Europe, September 2017 (URL: http://datadrivenadvertising.eu/wp-content/uploads/2017/09/EuropeOnline_FINAL.pdf), p. 7.

Boomers [born after 1969], and 8 out of 10 Pre-Boomers [born before 1969], distrust marketers and advertisers with their data".[12]

In 2016 a Eurobarometer survey of 26,526 people across the European Union found that:

> "Six in ten (60%) respondents have already changed the privacy settings on their Internet browser and four in ten (40%) avoid certain websites because they are worried their online activities are monitored. Over one third (37%) use software that protects them from seeing online adverts and more than a quarter (27%) use software that prevents their online activities from being monitored".[13]

This corresponds with an earlier Eurobarometer survey of similar scale in 2011, which found that "70% of Europeans are concerned that their personal data held by companies may be used for a purpose other than that for which it was collected".[14]

The same concerns arise in the United States. In May 2015, the Pew Research Centre reported that:

> "76% of [United States] adults say they are "not too confident" or "not at all confident" that records of their activity maintained by the online advertisers who place ads on the websites they visit will remain private and secure."[15]

In fact, respondents were the least confident in online advertising industry keeping personal data about them private than any other category of data processor, including social media platforms, search engines, and credit card companies. 50% said that no information should be shared with "online advertisers".[16]

---

[12] "GFK survey on data privacy and trust: data highlights", GFK, July 2015, p. 29.

[13] "Eurobarometer: e-Privacy (Eurobarometer 443)", European commission, December 2016 (URL: http://ec.europa.eu/COMMFrontOffice/publicopinion/index.cfm/Survey/getSurveyDetail/instruments/FLASH/surveyKy/2124), p. 5, 36-7.

[14] "Special Eurobarometer 359: attitudes on data protection and electronic identity in the European Union", European Commission, June 2011, p. 2.

[15] Mary Madden and Lee Rainie, "Americans' view about data collection and security", Pew Research Center, May 2015 (URL: http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/05/Privacy-and-Security-Attitudes-5.19.15_FINAL.pdf), p. 7.

[16] Mary Madden and Lee Rainie, "Americans' view about data collection and security", Pew Research Center, May 2015 (URL: http://assets.pewresearch.org/wp-content/uploads/sites/14/2015/05/Privacy-and-Security-Attitudes-5.19.15_FINAL.pdf), p. 25.

In a succession of surveys, large majorities express concern about ad tech. The UK's Royal Statistical Society published research on trust in data and attitudes toward data use and data sharing in 2014, and found that:

> "the public showed very little support for "online retailers looking at your past pages and sending you targeted advertisements", which 71% said should not happen".[17]

Similar results have appeared in the marketing industry's own research. RazorFish, an advertising agency, conducted a study of 1,500 people in the UK, US, China, and Brazil, in 2014 and found that 77% of respondents thought it was an invasion of privacy when advertising targeted them on mobile.[18]

These concerns are manifest in how people now behave online. The enormous growth of adblocking (to 615 million active devices by the start of 2017)[19] across the globe demonstrates the concern that Internet users have about being tracked and profiled by the ad tech industry companies. One industry commentator has called this the "biggest boycott in history".[20]

Concern about the misuse of personal data in online behavioural advertising is not confided to the public. Reputable advertisers, who pay for campaigns online, are concerned about it too. In January 2018, the CEO of the World Association of Advertisers, Stephan Loerke, wrote an opinion piece in AdAge attacking the current system as a "data free-for-all" where "each ad being served involved data that had been touched by up to fifty companies according to programmatic experts Labmatik".[21]

Correspondence with the industry on this matter to date

---

[17] "The data trust deficit: trust in data and attitudes toward data use and data sharing", Royal Statistical Society, July 2014, p. 5.

[18] Stephen Lepitak, "Three quarters of mobile users see targeted adverts as invasion of privacy, says Razorfish global research", The Drum, 30 June 2014 (URL: https://www.thedrum.com/news/2014/06/30/three-quarters-mobile-users-see-targeted-adverts-invasion-privacy-says-razorfish).

[19] "The state of the blocked web: 2017 global adblock report", PageFair, January 2017 (https://pagefair.com/downloads/2017/01/PageFair-2017-Adblock-Report.pdf).

[20] Doc Searls, "Beyond ad blocking – the biggest boycott in human history", Doc Searls Weblog, 28 September 2015 (https://blogs.harvard.edu/doc/2015/09/28/beyond-ad-blocking-the-biggest-boycott-in-human-history/).

[21] Stephan Loerke, "GDPR data-privacy rules signal a welcome revolution", AdAge, 25 January 2018 (URL: http://adage.com/article/cmo-strategy/gdpr-signals-a-revolution/312074/).

On 16 January 2018 I wrote to representatives of the IAB Europe working group (via IAB UK) to privately give feedback on a private draft of the IAB-led industry response to GDPR. I highlighted the following.

> First, bid requests would leak personal data among many parties without any protection. This would infringe Article 5 of the GDPR.
>
> Second, a lack of granularity and informed choice in the IAB's consent framework arose from the conflation of many separate purposes under a small number of nebulous purposes, and inadequate information. This would render consent invalid.

Although I was thanked for my input, I received no substantive response.

On 21 February 2018, in a video call, I raised concerns about the leakage of personal data in bid requests with the coordinator of the IAB TechLab working group responsible for designing an update to the new OpenRTB specification.

But when the IAB published its GDPR "framework" in March I learned that none of these concerns had been addressed. On 20 March 2018, I published my original feedback in an open letter. This is online at https://pagefair.com/blog/2018/iab-europe-consent-problems/.

On 4 September 2018 I wrote a detailed letter to the IAB and to IAB TechLab on behalf of Brave, to highlight critical data protection flaws in OpenRTB 3, an update to the RTB specification on which the IAB has solicited feedback. I set out in detail the acute hazard of broadcasting the personal data of a website visitor in bid requests, every time that the visitor loads a page. The letter I sent is available at https://brave.com/iab-rtb-problems/feedback-on-the-beta-OpenRTB-3.0-specification-.pdf.

On 5 September 2018, the IAB responded with a four line email that rejected the matter:

> **Feedback on the beta OpenRTB 3.0 specification**
>
> <*@iabtechlab.com> Wed, Sep 5, 2018 at 6:46 PM
> To: Johnny Ryan <*@brave.com>, OpenMedia <openmedia@iabtechlab.com>
> Cc: <*@iabtechlab.com>, <*@iabtechlab.com>
>
> Johnny,
>
> Thank you for submitting this feedback to the OpenRTB working group; your feedback has been shared with OpenRTB and Tech Lab leadership. It is (and always has been) the responsibility of

companies themselves to be aware of any and all relevant laws and regulations, and to adjust their platforms and practices to be compliant. In this case, any implementer of OpenRTB who should also be complying with GDPR could do so perhaps by using the Transparency and Consent Framework to communicate consumer consent and/or legitimate interest. OpenRTB represents protocol, not policy.

Thank you,
Jennifer & OpenRTB working group

Jennifer Derke
Director of Product, Automation/Programmatic
IAB Tech Lab
San Francisco, CA
[Quoted text hidden]

APPENDICES

Appendix 1. What personal data are shared in OpenRTB bid requests?

This summary list is incomplete. Other fields may contain personal data.[22]

"Site"[23]

- The specific URL that a visitor is loading, which shows what they are reading or watching.

"Device"[24]

- Operating system and version.
- Browser software and version.
- IP address.
- Device manufacturer, model, and version.
- Height, width, and ratio of screen.
- Whether JavaScript is supported.
- The version of Flash supported by the browser.
- Language settings.
- Carrier / ISP.
- Type of connection, if mobile.
- Network connection type.
- Hardware device ID (hashed).
- MAC address of the device (hashed).

"User"[25]

- An Ad Exchange's unique personal identifier for the visitor to the website. (This may rotate, but the specification says that it "must be stable long enough to serve reasonably as the basis for frequency capping and retargeting."[26])
- Advertiser's "buyeruid", a unique personal identifier for the data subject.
- The website visitor's year of birth, if known.
- The website visitor's gender, if known.
- The website visitor's interests.
- Additional data about the website visitor, if available from a data broker.[27] (These may include the "segment"[28] category previously decided by the data broker, based on the broker's previous profiling of this particular person.)

---

[22] For example, thirty eight of the data fields in the specification contain the phrase "optional vendor specific extensions".

[23] "Object: site" in "AdCOM Specification v1.0, Beta Draft", IAB TechLab, 24 July 2018 (URL: https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--site-).

[24] "Object: device" in ibid.

[25] "Object: device" in ibid.

[26] ibid.

[27] "Object: data" in ibid.

[28] "Object: segment" in ibid.

"Geo"[29]

- Location latitude and longitude.
- Zip/postal code.

[29] "Object: geo" in ibid.

Appendix 2. What personal data are shared in Google's proprietary bid requests?

"Publisher"[30]

- The specific URL that a visitor is loading, which shows what they are reading or watching. Note that sometimes publishers using Google's system prevent their URL from being shared.[31]

"Device"

- Operating system and version.
- Browser software and version (some data may be partially redacted).[32]
- Device manufacturer, model, and version.
- Height, width, and ratio of screen.
- Language settings.
- Carrier.
- Type of connection, if mobile.
- Hardware device IDs[33] (in "some circumstances", Google may impose "special constraints" on this. These constraints are not defined)[34]

"User"

- The Google ID of the website visitor
  (May be subject to some form of undefined "special constraints" in "some circumstances".)[35]
- Google's "Cookie Match Service" results, which enables a recipient to determine if the website visitor is a person they already have a profile of, and to combine their existing data with new data in the bid request.[36]

---

[30] All items in this appendix are drawn from "Authorized Buyers Real-Time Bidding Proto", Google, 5 September 2018 (URL: https://developers.google.com/authorized-buyers/rtb/realtime-bidding-guide).

[31] "Set your mobile app inventory to Anonymous or Branded in Ad Exchange", Google Ad Manager Help (URL: https://support.google.com/admanager/answer/6334919?hl=en)

[32] "Certain data may be redacted or replaced", see "user_agent" in "Authorized Buyers Real-Time Bidding Proto", Google, 5 September 2018 (URL: https://developers.google.com/authorized-buyers/rtb/realtime-bidding-guide).

[33] Some fields (such as advertising_id) are sent encrypted, but recipients can decrypt using keys that Google gives them when they set up their accounts, or are sent using standard encrypted SSL web connections. See "Decrypt Advertising ID", Authorized Buyers, Google (URL: https://developers.google.com/authorized-buyers/rtb/response-guide/decrypt-advertising-id).

[34] "In some circumstances there are special constraints on what can be done with user data for an ad request". Google vaguely states that in such a case, "user-related data will not be sent unfettered". User ID, Android or Apple device advertising ID, and "cookie match" data can be affected. See "User Data Treatments", Authorized Buyers, Google (URL: https://developers.google.com/authorized-buyers/rtb/user_data_treatments).

[35] ibid.

[36] "Cookie Matching", Google, 5 September 2018 (URL: https://developers.google.com/authorized-buyers/rtb/cookie-guide?hl=en).

(May be subject to some form of undefined "special constraints" in "some circumstances".)[37]

- The website visitor's interests.
- Whether the website visitor is present on a particular "user list" of targeted people (which may be a category previously decided by an advertiser, or the data broker they acquired the data from, based on the broker's previous profiling of this particular person).

"Location"

- Location latitude and longitude.
- Zip/postal code, or postal code prefix if a full post code is unavailable.
- Whether the user is present within a small "hyper local" area.

[37] see note 36.

Appendix 3. Selected data tables from OpenRTB bid request specification documents

The following tables are copied from AdCOM specification v1, which is part of the OpenRTB 3.0 specification.[38] This defines what data can be included in a bid request. Only selected tables relevant to website bid requests are included here. URLs of the specific part of the specification from where the tables are taken are presented above each table.

Publisher

Object: Site

Derived from: DistributionChannel

This object is used to define an ad supported website, in contrast to a non-browser application, for example. As a derived class, a "Site" object inherits all "DistributionChannel" attributes and adds those defined below.

Attribute	Type	Definition
domain	string	Domain of the site (e.g., "mysite.foo.com").
cat	string array	Array of content categories describing the site using IDs from the taxonomy indicated in "cattax".
sectcat	string array	Array of content categories describing the current section of the site using IDs from the taxonomy indicated in "cattax".
pagecat	string array	Array of content categories describing the current page or view of the site using IDs from the taxonomy indicated in "cattax".
cattax	integer	The taxonomy in use for the "cat", "sectcat" and "pagecat" attributes. Refer to List: Category Taxonomies.
privpolicy	integer	Indicates if the site has a privacy policy, where 0 = no, 1 = yes.
keywords	string	Comma separated list of keywords about the site.
page	string	URL of the page within the site.
ref	string	Referrer URL that caused navigation to the current page.
search	string	Search string that caused navigation to the current page.
mobile	integer	Indicates if the site has been programmed to optimize layout when viewed on mobile devices, where 0 = no, 1 = yes.
amp	integer	Indicates if the page is built with AMP HTML, where 0 = no, 1 = yes.
ext	object	Optional vendor-specific extensions.

https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--site-

38 "AdCOM Specification v1.0, Beta Draft", IAB TechLab, 24 July 2018 (URL: https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md).

Object: Publisher

This object describes the publisher of the media in which ads will be displayed.

Attribute	Type	Definition
id	string, recommended	Vendor-specific unique publisher identifier, as used in ads.txt files.
name	string	Displayable name of the publisher.
domain	string	Highest level domain of the publisher (e.g., "publisher.com").
cat	string array	Array of content categories that describe the publisher using IDs from the taxonomy indicated in "cattax".
cattax	integer	The taxonomy in use for the "cat" attribute. Refer to List: Category Taxonomies.
ext	object	Optional vendor-specific extensions.

https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--publisher-

User

Object: User

This object contains information known or derived about the human user of the device (i.e., the audience for advertising). The user ID is a vendor-specific artifact and may be subject to rotation or other privacy policies. However, this user ID must be stable long enough to serve reasonably as the basis for frequency capping and retargeting.

Attribute	Type	Definition
id	string; recommended	Vendor-specific ID for the user. At least one of "id" or "buyeruid" is strongly recommended.
buyeruid	string; recommended	Buyer-specific ID for the user as mapped by an exchange for the buyer. At least one of "id" or "buyeruid" is strongly recommended.
yob	integer	Year of birth as a 4-digit integer.
gender	string	Gender, where "M" = male, "F" = female, "O" = known to be other (i.e., omitted is unknown).
keywords	string	Comma separated list of keywords, interests, or intent.
consent	string	GDPR consent string if applicable, complying with the comply with the IAB standard Consent String Format in the Transparency and Consent Framework technical specifications.
geo	object	Location of the user's home base (i.e., not necessarily their current location). Refer to Object: Geo.
data	object array	Additional user data. Each "Data" object represents a different data source. Refer to Object: Data.
ext	object	Optional vendor-specific extensions.

https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--user-

Object: Data

The data and segment objects together allow additional data about the related object (e.g., user, content) to be specified. This data may be from multiple sources whether from the exchange itself or third parties as specified by the "id" attribute. When in use, vendor-specific IDs should be communicated *a priori* among the parties.

Attribute	Type	Definition
id	string	Vendor-specific ID for the data provider.
name	string	Vendor-specific displayable name for the data provider.
segment	object array	Array of "Segment" objects that contain the actual data values. Refer to Object: Segment.
ext	object	Optional vendor-specific extensions.

https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--data-

Object: Segment

Segment objects are essentially key-value pairs that convey specific units of data. The parent "Data" object is a collection of such values from a given data provider. When in use, vendor-specific IDs should be communicated *a priori* among the parties.

Attribute	Type	Definition
id	string	ID of the data segment specific to the data provider.
name	string	Displayable name of the data segment specific to the data provider.
value	string	String representation of the data segment value.
ext	object	Optional vendor-specific extensions.

https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--segment-

Device

Object: Device

This object provides information pertaining to the device through which the user is interacting. Device information includes its hardware, platform, location, and carrier data. The device can refer to a mobile handset, a desktop computer, set top box, or other digital device.

Attribute	Type	Definition
type	integer	The general type of device. Refer to List: Device Types.
ua	string	Browser user agent string.
ifa	string	ID sanctioned for advertiser use in the clear (i.e., not hashed).
dnt	integer	Standard "Do Not Track" flag as set in the header by the browser, where 0 = tracking is unrestricted, 1 = do not track.
lmt	integer	"Limit Ad Tracking" signal commercially endorsed (e.g., iOS, Android), where 0 = tracking is unrestricted, 1 = tracking must be limited per commercial guidelines.
make	string	Device make (e.g., "Apple").
model	string	Device model (e.g., "iPhone").
os	integer	Device operating system. Refer to List: Operating Systems.
osv	string	Device operating system version (e.g., "3.1.2").
hwv	string	Hardware version of the device (e.g., "5S" for iPhone 5S).
h	integer	Physical height of the screen in pixels.
w	integer	Physical width of the screen in pixels.
ppi	integer	Screen size as pixels per linear inch.
pxratio	float	The ratio of physical pixels to device independent pixels.
js	integer	Support for JavaScript, where 0 = no, 1 = yes.
lang	string	Browser language using ISO-639-1-alpha-2.
ip	string	IPv4 address closest to device.
ipv6	string	IP address closest to device as IPv6.
xff	string	The value of the x-forwarded-for header.
iptr	integer	Indicator of truncation of any of the IP attributes (i.e., "ip", "ipv6", "xff"), where 0 = no, 1 = yes (e.g., from 1.2.3.4 to 1.2.3.0). Refer to tools.ietf.org/html/rfc6235#section-4.1.1 for more information on IP truncation.
carrier	string	Carrier or ISP (e.g., "VERIZON") using exchange curated string names which should be published to bidders a priori.
mccmnc	string	Mobile carrier as the concatenated MCC-MNC code (e.g., "310-005" identifies Verizon Wireless CDMA in the USA). Refer to en.wikipedia.org/wiki/Mobile_country_code for further information and references. Note that the dash between the MCC and MNC parts is required to remove parsing ambiguity.
mccmncsim	string	MCC and MNC of the SIM card using the same format as "mccmnc". When both values are available, a difference between them reveals that a user is roaming.
contype	integer	Network connection type. Refer to List: Connection Types.
geofetch	integer	Indicates if the geolocation API will be available to JavaScript code running in display ad,

geofetch	integer	Indicates if the geolocation API will be available to JavaScript code running in display ad, where 0 = no, 1 = yes.
geo	object	Location of the device (i.e., typically the user's current location). Refer to Object: Geo.
ext	object	Optional vendor-specific extensions.

https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--device-

Location

Object: Geo

This object encapsulates various methods for specifying a geographic location. When subordinate to a "Device" object, it indicates the location of the device which can also be interpreted as the user's current location. When subordinate to a "User" object, it indicates the location of the user's home base (i.e., not necessarily their current location).

The "lat" and "lon" attributes should only be passed if they conform to the accuracy depicted in the "type" attribute. For example, the centroid of a large region (e.g., postal code) should not be passed.

Attribute	Type	Definition
type	integer	Source of location data; recommended when passing lat/lon. Refer to List: Location Types.
lat	float	Latitude from -90.0 to +90.0, where negative is south.
lon	float	Longitude from -180.0 to +180.0, where negative is west.
accur	integer	Estimated location accuracy in meters; recommended when lat/lon are specified and derived from a device's location services (i.e., type = 1). Note that this is the accuracy as reported from the device. Consult OS specific documentation (e.g., Android, iOS) for exact interpretation.
lastfix	integer	Number of seconds since this geolocation fix was established. Note that devices may cache location data across multiple fetches. Ideally, this value should be from the time the actual fix was taken.
ipserv	integer	Service or provider used to determine geolocation from IP address if applicable (i.e., "type" = 2). Refer to List: IP Location Services.
country	string	Country code using ISO-3166-1-alpha-2. Note that alpha-3 codes may be encountered and vendors are encouraged to be tolerant of them.
region	string	Region code using ISO-3166-2; 2-letter state code if USA.
metro	string	Regional marketing areas such as Nielsen's DMA codes or other similar taxonomy to be agreed among vendors prior to use. Note that DMA is a trademarked asset of The Nielsen Company. Vendors are encouraged to ensure their use of DMAs is properly licensed.
city	string	City using United Nations Code for Trade & Transport Locations "UN/LOCODE" with the space between country and city suppressed (e.g., Boston MA, USA = "USBOS"). Refer to UN/LOCODE Code List.
zip	string	ZIP or postal code.
utcoffset	integer	Local time as the number +/- of minutes from UTC.
ext	object	Optional vendor-specific extensions.

https://github.com/InteractiveAdvertisingBureau/AdCOM/blob/master/AdCOM%20BETA%201.0.md#object--geo-

Appendix 4. Selected data tables from Google ("Authorised Buyer") RTB bid request specification documents

The following tables are copied from Google's RTB documentation.[39] This defines what data can be included in a bid request. Only selected tables relevant to website bid requests are included here. URLs of the specific part of the specification from where the tables are taken are presented above each table.

[39] "Authorized Buyers Real-Time Bidding Proto", Google, 5 September 2018 (URL: https://developers.google.com/authorized-buyers/rtb/realtime-bidding-guide)

User

google_user_id	optional	string	The Google ID for the user as described in the documentation for the cookie matching service. This field is the unpadded web-safe base64 encoded version of a binary cookie ID. See the Base 64 Encoding with URL and Filename Safe Alphabet section in RFC 3548 for encoding details This field is the same as the Google ID returned by the cookie matching service. Not set if there is one or more **user_data_treatment** value, see **constrained_usage_google_user_id** instead.
constrained_usage_google_user_id	optional	string	Only set if there is one or more **user_data_treatment** value. If **constrained_usage_google_user_id** is set, then **google_user_id** is not set. You must be whitelisted for all **user_data_treatments** in this request in order to receive this field.
cookie_version	optional	uint32	The version number of the **google_user_id**. We may sometimes change the mapping from cookie to **google_user_id**. In this case the version will be incremented.
cookie_age_seconds	optional	int32	The time in seconds since the **google_user_id** was created. This number may be quantized.
hosted_match_data	optional	bytes	Match data stored for this **google_user_id** through the cookie matching service. If a match exists, then this field holds the decoded data that was passed in the **google_hm** parameter. Not set if there is one or more **user_data_treatment** value, see **constrained_usage_hosted_match_data** instead.
constrained_usage_hosted_match_data	optional	bytes	Only set if there is one or more **user_data_treatment** value. If **constrained_usage_hosted_match_data** is set, then **hosted_match_data** is not set. You must be whitelisted for all **user_data_treatments** in this request in order to receive this field.
user_agent	optional	string	A string that identifies the browser and type of device that sent the request. Certain data may be redacted or replaced.
publisher_country	optional	string	The billing address country of the publisher. This may be different from the detected country of the user in **geo_criteria_id** or the hosting country of the website. For a complete list of country codes, see the country codes list in the AdWords API documentation.
geo_criteria_id	optional	int32	Location of the end user. Uses a subset of the

			API documentation.
geo_criteria_id	optional	int32	Location of the end user. Uses a subset of the codes used in the AdWords API. See the geo-table.csv table in the technical documentation for list of IDs. The **geo_criteria_id** field replaces the deprecated country, region, city, and metro fields.
postal_code **postal_code_prefix**	optional	string	Detected postal code of the appropriate type for the country of the end user (e.g., zip code if the country is "US"). The **postal_code_prefix** field is set when accuracy is too low to imply a full code otherwise the **postal_code** field is set.
encrypted_hyperlocal_set	optional	bytes	Hyperlocal targeting signal when available, encrypted as described in the Decrypt Hyperlocal Target Signals guide.
hyperlocal_set	optional	HyperlocalSet	Unencrypted version of **encrypted_hyperlocal_set**. This field is only set when using an SSL connection.
timezone_offset	optional	int32	The offset of the user's time from GMT in minutes. For example, GMT+10 is **timezone_offset = 600**.
user_vertical	repeated	int32	List of detected user verticals. Currently unused.
user_list	repeated	UserList	

UserList object

This field is not populated by default. We recommend that bidders instead store and look up list IDs using either `google_user_id` or `hosted_match_data` as keys.

Attribute	Required/Optional	Type	Implementation details
id	optional	int64	The user list ID.
age_seconds	optional	int32	The time in seconds since the user was added to the list.

advertising_id	optional	bytes	Unencrypted version of **encrypted_advertising_** This field is only set when using an SSL connection. T field is a 16 byte UUID.
hashed_idfa	optional	bytes	Unencrypted version of **encrypted_hashed_idfa**. This field is only set when using an SSL connection. T field is a 16 byte MD5.
constrained_usage_ encrypted_advertising_id	optional	bytes	Only set if the BidRequest contains one or more **user_data_treatment** value. If **constrained_usage_encrypted_advertising** or **constrained_usage_encrypted_hashed_id** is set, then the corresponding non-constrained field is set. You must be whitelisted for all **user_data_treatments** in this request in order to receive these fields.
constrained_usage_advertising_id	optional	bytes	Unencrypted version of **constrained_usage_encrypted_advertising** This field is only set when using an SSL connection. T field is a 16 byte UUID.
constrained_usage_ encrypted_hashed_idfa	optional	bytes	
constrained_usage_hashed_idfa	optional	bytes	Unencrypted version of **constrained_usage_encrypted_hashed_idfa** This field is only set when using an SSL connection. T field is a 16 byte MD5.
app_name	optional	string	App names for Android apps are from the Google Pla store. App names for iOS apps are provided by App Annie.
app_rating	optional	float	Average user rating for the app. The range of user rati is between 1.0 and 5.0. Currently only available for ap in Google Play store.

Mobile object

Information for ad queries coming from mobile devices. A mobile device is either a smartphone or a tablet. This is present for ad queries both from mobile devices browsing the web and from mobile apps.

Attribute	Required/Optional	Type	Implementation details
is_app	optional	bool	If true, then this request is from a mobile application. always be true when **app_id** is set. May also be true anonymous inventory, in which case **anonymous_id** be set.
app_id	optional	string	The identifier of the mobile app when this ad query comes from a mobile app. If the app was downloaded from the Apple iTunes app store, then this is the app-store ID, e.g., 343200656. For Android devices, this is fully qualified package name, e.g., com.rovio.angrybir For Windows devices it's the App ID, e.g., f15abcde-f6 47i0-j3k8-37l93817mn3o.
is_interstitial_request	optional	bool	If true, then this is a mobile full screen ad request.
app_category_ids	repeated	int32	This field contains the IDs of categories to which the current mobile app belongs. This field will be empty if **is_app** is false. The mapping between mobile apps a categories is defined by the Google Play Store for Android apps, or the Apple iTunes Store for iOS apps. look up category name from category ID, refer to the mobile app categories table.
is_mobile_web_optimized	optional	bool	For a mobile web request, this field indicates whether page is optimized for mobile browsers on high-end mobile phones. default=false
encrypted_advertising_id	optional	bytes	This field is used for advertising identifiers for: 1) iOS devices (This is called Identifier for Advertising IDFA, as described in this Help Center article.) 2) Android devices. 3) Roku devices. 4) Microsoft Xbox devices. 5) Amazon devices. When the **encrypted_advertising_id** is an IDFA plaintext after decrypting the ciphertext is the IDFA (1 byte UUID) returned by iOS's **[ASIdentifierManag advertisingIdentifier]**. For **encrypted_hashed_idfa**, the plaintext is the 16 by MD5 hash of the IDFA. Only one of the two fields will b available, depending on the version of the SDK makin the request. Later SDKs provide unhashed values. Th are not set if there is one or more **user_data_treatment** value in the BidRequest, se **constrained_usage_encrypted_advertising** and **constrained_usage_encrypted_hashed_i** instead.
encrypted_hashed_idfa	optional	bytes	See also the description for **encrypted_advertising_id**.
advertising_id	optional	bytes	Unencrypted version of **encrypted_advertising_** This field is only set when using an SSL connection. T

advertising_id	optional	bytes	Unencrypted version of **encrypted_advertising_** This field is only set when using an SSL connection. T field is a 16 byte UUID.
hashed_idfa	optional	bytes	Unencrypted version of **encrypted_hashed_idfa**. This field is only set when using an SSL connection. T field is a 16 byte MD5.
constrained_usage_ encrypted_advertising_id	optional	bytes	Only set if the BidRequest contains one or more **user_data_treatment** value. If **constrained_usage_encrypted_advertising** or **constrained_usage_encrypted_hashed_id** is set, then the corresponding non-constrained field is set. You must be whitelisted for all **user_data_treatments** in this request in order to receive these fields.
constrained_usage_advertising_id	optional	bytes	Unencrypted version of **constrained_usage_encrypted_advertising** This field is only set when using an SSL connection. T field is a 16 byte UUID.
constrained_usage_ encrypted_hashed_idfa	optional	bytes	
constrained_usage_hashed_idfa	optional	bytes	Unencrypted version of **constrained_usage_encrypted_hashed_idfa**. This field is only set when using an SSL connection. T field is a 16 byte MD5.
app_name	optional	string	App names for Android apps are from the Google Pla store. App names for iOS apps are provided by App Annie.
app_rating	optional	float	Average user rating for the app. The range of user rati is between 1.0 and 5.0. Currently only available for ap in Google Play store.

Publisher

This section lists information that we know about the web page or mobile application where the impression originates.

Attribute	Required/Optional	Type	Implementation details
publisher_id	optional	string	The publisher ID as defined by the publisher code suffix of the web property code. For instance, "pub-123" is the publisher code of web property code "ca-pub-123" (ca- is the product specific prefix of the web property).
seller_network_id	optional	int32	The seller network ID. See seller-network-ids.txt file in the technical documentation for a list of IDs. This is only set if the site is not anonymous and the publisher allows site targeting.
partner_id	optional	fixed64	ID for the partner that provides this inventory. This is only set when **seller_network_id** is also set and further partner information beyond the **seller_network_id** is also available. The value of the **partner_id** is not meaningful beyond providing a stable identifier.
url	optional	string	The URL of the page with parameters removed. This is only set if the site is not anonymous and the publisher allows site targeting. You can use **anonymous_id** for targeting if the inventory is anonymous. Otherwise, use **detected_vertical**. Only one of **url** or **anonymous_id** is ever set in the same request. This always starts with a protocol (either http or https).
anonymous_id	optional	string	An id for the domain of the page. This is set when the inventory is anonymous. Only one of **url** or **anonymous_id** is ever set in the same request.
detected_language	repeated	string	Detected user languages, based on the language of the web page, the browser settings, and other signals. The order is arbitrary. The codes are 2 or 5 characters and are documented in the language codes table.
detected_vertical	repeated	Vertical	Unordered list of detected content verticals. See the publisher-verticals.txt file in the technical documentation for a list of IDs.
detected_content_label	repeated	int32	List of detected content labels. See the content-labels.txt file in the technical documentation for a list of IDs.
device	optional	Device	

device	optional	Device	
key_value	repeated	KeyValue	
mobile	optional	Mobile	
video	optional	Video	
publisher_settings_list_id	optional	fixed64	The publisher settings list ID that applies to this page. See the RTB Publisher Settings guide for details.
publisher_type	optional	PublisherType	Publisher type of the inventory where the ad will be shown. For an Authorized Buyers publisher, its inventory can be either owned and operated (O&O), represented by the publisher, or of unknown status. AdSense and AdMob inventory is represented by Google. **enum PublisherType** **UNKNOWN_PUBLISHER_TYPE = 0;** **ADX_PUBLISHER_OWNED_AND_OPERATED = 1;** **ADX_PUBLISHER_REPRESENTED = 2;** **GOOGLE_REPRESENTED = 3;** **default = UNKNOWN_PUBLISHER_TYPE**
adslot	repeated	AdSlot	
bid_response_feedback	repeated	BidResponseFeedback	

Vertical object

One or more detected verticals for the page as determined by Google.

Attribute	Required/Optional	Type	Implementation details
id	required	int32	The vertical ID. See the publisher-verticals.txt file in the technical documentation for a list of IDs.
weight	required	float	Weight for this vertical, in the (0.0, 1.0] range. More relevant verticals have higher weights.

Location

Hyperlocal object

A hyperlocal targeting location when available.

Attribute	Required/Optional	Type	Implementation details
`corners`	repeated	Point	The mobile device can be at any point inside the geofence polygon defined by a list of corners. Currently, the polygon is always a parallelogram with 4 corners.

Point object

A location on the Earth's surface.

Attribute	Required/Optional	Type	Implementation details
`latitude`	optional	float	Latitude of the location.
`longitude`	optional	float	Longitude of the location.

HyperlocalSet object

Attribute	Required/Optional	Type	Implementation details
`hyperlocal`	repeated	Hyperlocal	This field currently contains at most one hyperlocal polygon.
`center_point`	optional	Hyperlocal.Point	The approximate geometric center of the geofence area. It is calculated exclusively based on the geometric shape of the geofence area and in no way indicates the mobile device's actual location within the geofence area. If multiple hyperlocal polygons are specified above then **`center_point`** is the geometric center of all hyperlocal polygons.
`encrypted_hyperlocal_set`	optional	bytes	Hyperlocal targeting signal when available, encrypted as described in this guide

Device

Device object

Information about the device.

Attribute	Required/Optional	Type	Implementation details
DeviceType		enum	**UNKNOWN_DEVICE = 0;** **HIGHEND_PHONE = 1;** **TABLET = 2;** **PERSONAL_COMPUTER = 3;** - Desktop or laptop devices. **CONNECTED_TV = 4;** - Both connected TVs (smart TVs) and connected devices (such as Roku and Apple TV). **GAME_CONSOLE = 5;**
device_type	optional	DeviceType	default = **UNKNOWN_DEVICE**
platform	optional	string	The platform of the device. Examples: Android, iPhone, Palm
brand	optional	string	The brand of the device, e.g., Nokia, Samsung.
model	optional	string	The model of the device, e.g., N70, Galaxy.
os_version	optional	OsVersion	The OS version; e.g., 2 for Android 2.1, or 3.3 for iOS 3.3.1.
carrier_id	optional	int64	Unique identifier for the mobile carrier if the device is connected to the internet via a carrier (as opposed to via WiFi). To look up carrier name and country from carrier ID, refer to this mobile carriers table.
screen_width	optional	int32	The width of the device screen in pixels.
screen_height	optional	int32	The height of the device screen in pixels.
screen_pixel_ratio_millis	optional	int32	Used for high-density devices (e.g., iOS retina displays). A non-default value indicates that the nominal screen size (with pixels as the unit) does not describe the actual number of pixels in the screen. For example, nominal width and height may be 320x640 for a screen that actually has 640x1080 pixels, in which case **screen_width=320**, **screen_height=640**, and **screen_pixel_ratio_millis=2000**, since each axis has twice as many pixels as its dimensions would indicate. default = 0
screen_orientation	optional	ScreenOrientation	The screen orientation of the device when the ad request is sent. enum **ScreenOrientation** **UNKNOWN_ORIENTATION = 0;** **PORTRAIT = 1;** **LANDSCAPE = 2;** **default = UNKNOWN_ORIENTATION**
hardware_version	optional	string	Apple iOS device model, e.g., "iphone 5s", "iphone 6+", "ipad 4".

OSVersion object

Contains the OS version of the platform. For instance, for Android 2, major=2, minor=0. For iPhone 3.3.1, major=3 and minor=3.

Attribute	Required/Optional	Type
major **minor** **micro**	optional	int32