Privacy, Security, and Network Effects

Dr. Rob van Eijk
Managing Director for Europe, Future of Privacy Forum
Director Blaeu Privacy Response Team B.V.

FUTURE OF PRIVACY FORUM

5 February 2020 | I-Interim Rijk Data Science Crash Course | The Hague

Privacy

5 February 2020 | I-Interim Rijk Data Science Crash Course | The Hague

FUTURE OF PRIVACY FORUM

A definition for privacy

„Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated.”

[Westin, 1967, p. 7]

Question: what does privacy mean nowadays to you?

Foundations for the right to privacy in Europe [1]

Article 16 of the **Treaty on the Functioning of the European Union (TFEU)** which provides the legal basis for the adoption of Union legal instruments relating to the protection of personal data.

Foundations for the right to privacy in Europe [2]

- Article 7 and Article 8 of the **Charter of Fundamental Rights of the European Union (CFREU)**, the charter has a similar legal value as the TFEU.
- **Article 7**: „*Everyone has the right to respect for his or her private and family life, home and communications.*"

Foundations for the right to privacy in Europe [3]

Article 8 sees to the protection of personal data:

„(1) Everyone has the right to the protection of personal data concerning him or her,

(2) such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law. Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified,

(3) compliance with these rules shall be subject to control by an independent authority."

Foundations for the right to privacy in Europe [4]

- **General Data Protection Regulation (EU) 2016/679 (GDPR)** repealing General Data Protection Directive 95/46/EC (GDPD).
- The application and interpretation of the legal norms in the GDPR **conform to the CFREU**.

Eight OECD principles (1980, 2013)

- Collection limitation
- Data quality
- Purpose specification
- Use limitation
- Security safeguards
- Openness
- Individual participation, and
- Accountability

The GDPR aligns with the OECD principles

- Lawfulness, fairness, and transparency (Article 5(1a), Art. 6 GDPR)
- Purpose limitation (Article 5(1b) GDPR)
- Data minimization (Article 5(1c) GDPR)
- Accuracy (Article 5(1d) GDPR)
- Storage limitation (Article 5(1e) GDPR)
- Integrity and confidentiality (Article 5(1f) GDPR), and
- Accountability (Article 5(2) GDPR)

Privacy and other fundamental rights

- Freedom of expression, religion, and assembly/association
- Non-discrimination
- Presumption of innocence/right of defense
- Freedom of speech, freedom of thought, freedom of movement, and
- Right to liberty

Data protection in court

- EU law supersedes national law.
- A judge will apply a legal norm against the background of the aforementioned principles and fundamental rights.

SyRI legislation in violation of the right to privacy

- ECLI:NL:RBDHA:2020:865
- Cf. Article 8, paragraph 2 of the European Convention on Human Rights
- “The Court is of the opinion that the SyRI legislation does not provide sufficient safeguards to protect the right to respect for private life in relation to the risk indicators and the risk model that can be implemented in a specific SyRI project.”

Question: personal data?

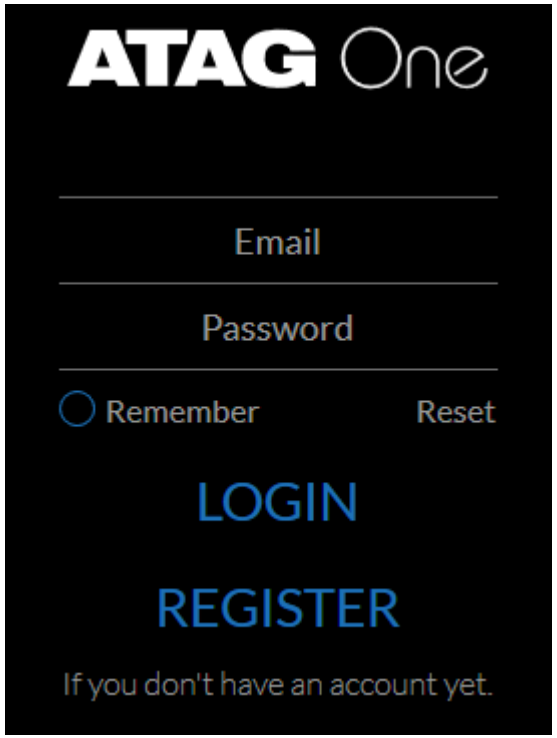
Meta data about my behavior

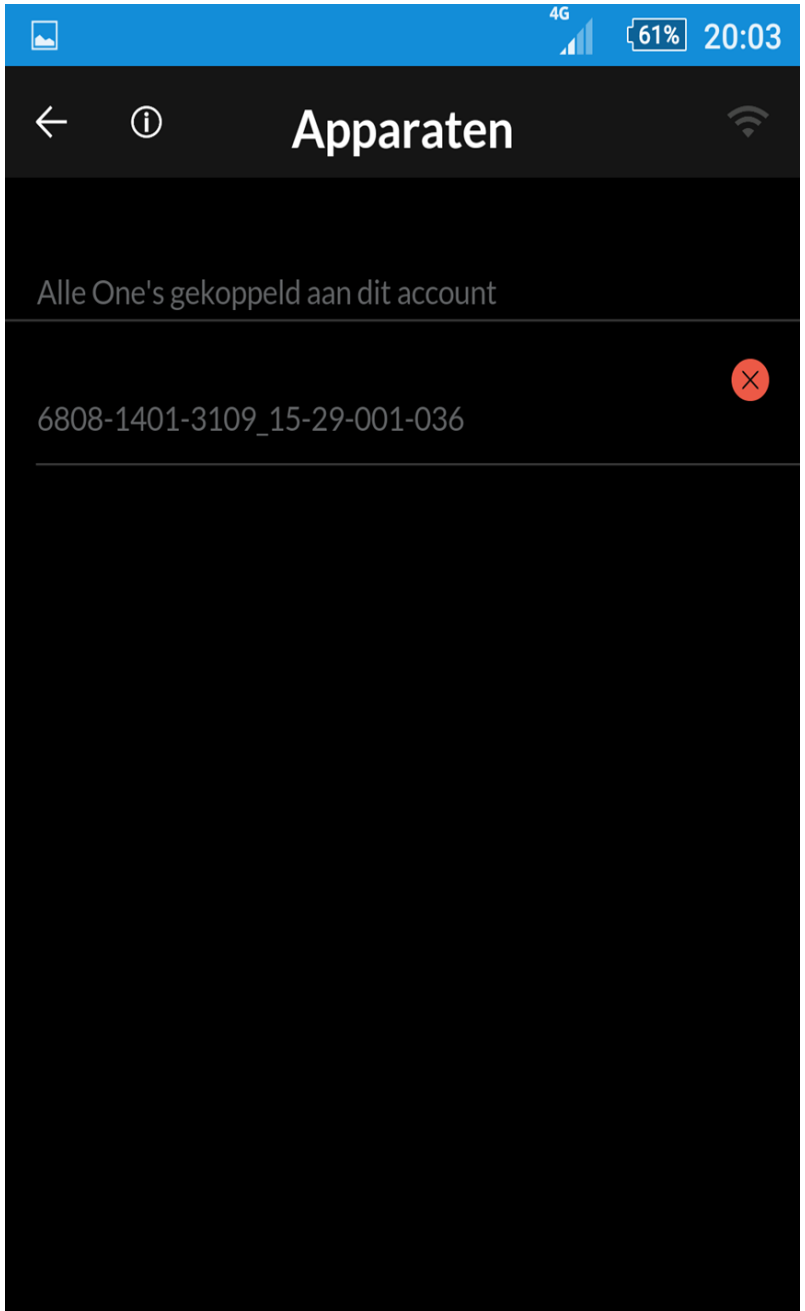
Device data:

- Type
- Heating capacity

Operational data:

- Date, time
- Burning hours
- Burning status
- Room temperature





Personal Data [1]

Article 4(1) GDPR:

Any information relating to an identified or identifiable natural person ('data subject’);

But… what is an identifiable natural person?

Personal Data [2]

Article 4(1) GDPR:

An identifiable natural person is one who can be identified, **directly or indirectly**,

in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or

to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

Personal Data [3]

Recital 26 GDPR: To determine whether a natural person is identifiable, account should be taken of **all the means reasonably likely to be used**, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly.

Note: in the online context personal data is often data about your behavior, i.e., (1) ***what*** you use and (2) ***how*** you use it.

Security

5 February 2020 | I-Interim Rijk Data Science Crash Course | The Hague

FUTURE OF PRIVACY FORUM

Privacy: a risk based approach [1]

(Recital 75 GDPR)

- The risk to the **rights and freedoms of natural persons**,
- of varying likelihood and severity, may result from personal data processing which could lead to
  - physical,
  - material or
  - non-material damage, in particular:

Privacy: a risk based approach [2]

(Recital 75 GDPR)

- where the processing may give rise to
  - discrimination,
  - identity theft or fraud,
  - financial loss,
  - damage to the reputation,
  - loss of confidentiality of personal data protected by professional secrecy,
  - unauthorized reversal of pseudonymization, or
  - any other significant economic or social disadvantage;

Privacy: a risk based approach [3]

(Recital 75 GDPR)

- where data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data;
- where personal data are processed which reveal **racial or ethnic origin, political opinions, religion or philosophical beliefs, trade union membership**,
- and the processing of **genetic data, data concerning health or data concerning sex life or criminal convictions and offences or related security measures**;

Privacy: a risk based approach [4]

(Recital 75 GDPR)

- where personal aspects are evaluated, in particular analyzing or predicting aspects concerning **performance at work, economic situation, health, personal preferences or interests, reliability or behavior, location or movements**, in order to create or use personal **profiles**;
- where personal data of vulnerable natural persons, in particular of **children**, are processed; or
- where processing involves a **large amount** of personal data and affects a **large number** of data subjects.

Processing [1]

(Article 4(2) GDPR)

Any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means.

Processing [2]

(Article 4(2) GDPR)

Such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

Pseudonymization [1]

(Article 4(5) GDPR)

- the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject **without the use of additional information**,
- provided that such additional information is kept separately and is subject to technical and organizational measures **to ensure that the personal data are not attributed to an identified or identifiable natural person**.

Pseudonymisation [2]

(Recital 26 GDPR)

Personal data which have undergone pseudonymization, which could be attributed to a natural person by the use of additional information **should be considered to be information on an identifiable natural person**.

Anonymization [1]

(Recital 26 GDPR)

- The principles of data protection should therefore **not apply to anonymous information**.
- Namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous **in such a manner that the data subject is not or no longer identifiable**.
- This Regulation does not therefore concern the processing of such anonymous information, **including for statistical or research purposes**.

Anonymization [2]

WP29 Opinion 5/2014 on Anonymization Techniques:

- Anonymization constitutes **a further processing of personal data**; as such, it must satisfy the requirement of compatibility by having regard to the legal grounds and circumstances of the further processing.

Anonymization [3]

WP29 Opinion 5/2014 on Anonymization Techniques :

The opinion elaborates on the robustness of each technique based on three criteria:

- is it still possible to **single out** an individual,
- is it still possible to **link records relating to an individual**, and
- can information be **inferred concerning an individual**?

Discussion [1]

Should a **WiFi MAC-address** be considered as **personal data?**

… and **in combination with additional data**, e.g.,

- the WiFi signal strength,
- time/date of the measurement,
- the location of a WiFi sensor?

Discussion [2]

What is your opinion on:

- hashing **on the sensor**,
- limiting measurements **in space and time** to specific times and locations?
- aggregating of WiFi data **on the server** instead of the WiFi sensor?

Do’s and don’ts

- Researchers Find 'Anonymized' Data Is Even Less Anonymous Than We Thought. URL: https://www.vice.com/en_ca/article/dygy8k/researchers-find-anonymized-data-is-even-less-anonymous-than-we-thought
- Privacy by default (versus by design)
- Pivacy by design
  - Select before you collect (data minimization)
  - Anonymization
  - Pseudonymization, privacy preserving data sharing
  - Encryption
  - Granular (consent) controls and revocation of consent
  - Data retention, persistency of identifiers
  - Data subject’s rights (DSARS)
  - Data portability
- Aim for surprise minimization !

Network Effects

5 February 2020 | I-Interim Rijk Data Science Crash Course | The hague

FUTURE OF PRIVACY FORUM

What is your view on cookies?

- *"I consent to the placement of cookies"*
- *"Before I consent, I read which cookies are placed"*
- *"Before I consent, I try to change the cookie settings"*
- *"I regularly delete my cookies"*
- *"I have a tool to protect my online privacy"*
- *"I sometimes refuse cookies"*
- *"I no longer visit a website because it places cookies"*

Freshening up - what are cookies?

- **Cookie:** text file saved and read on the peripherals
- An example cookie: NL 123456789 B01
- A cookie only becomes valuable after reading

NAME	s_sq
VALUE	%5B%5BB%5D%5D
DOMAIN	abnamro.nl
PATH	/
EXPIRES	At the end of the Session

NAME	s_cc
VALUE	true
DOMAIN	abnamro.nl
PATH	/
EXPIRES	At the end of the Session

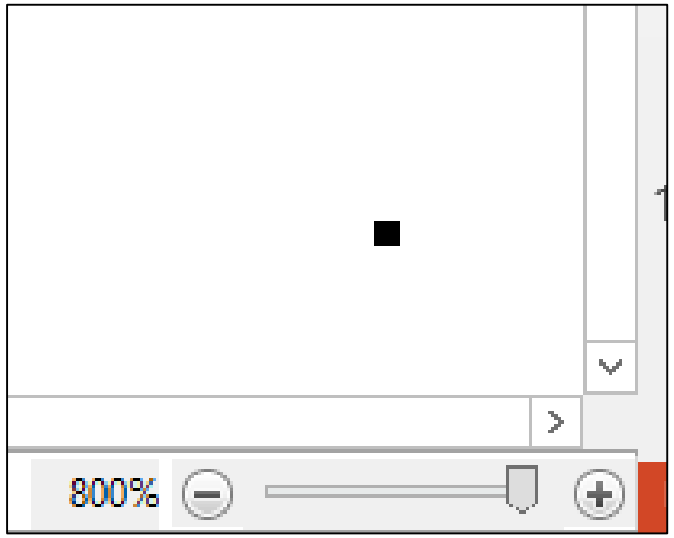
NAME	logonStateCookie
VALUE	I
DOMAIN	abnamro.nl
PATH	/
EXPIRES	At the end of the Session

Not just 'normal' cookies

JavaScript cookies: unieke nummers in de HTTP header of in de URL van een pagina (geen tekstbestandjes).

```
GET /pagead/id HTTP/1.1
Host: googleads.g.doubleclick.net
User-Agent: Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:67.0) Gecko/20100101 Firefox/67.0
Accept: */*
Accept-Language: en-US,en;q=0.5
Accept-Encoding: gzip, deflate, br
Origin: https://www.youtube.com
Connection: keep-alive
Referer:
https://www.youtube.com/embed/GSw2Ka7bxuI?autoplay=1&controls=0&disablekb=1&hl=nl&modestbranding=1&iv_load_policy=3
Cookie: IDE=AHWqTUmp5lzoZQOplvf88LFGS3StkyfK8tJc3QQlf90Hi1bqorkj5N7ex7Mrxh1f
Pragma: no-cache
Cache-Control: no-cache
```

Cookies and similar techniques

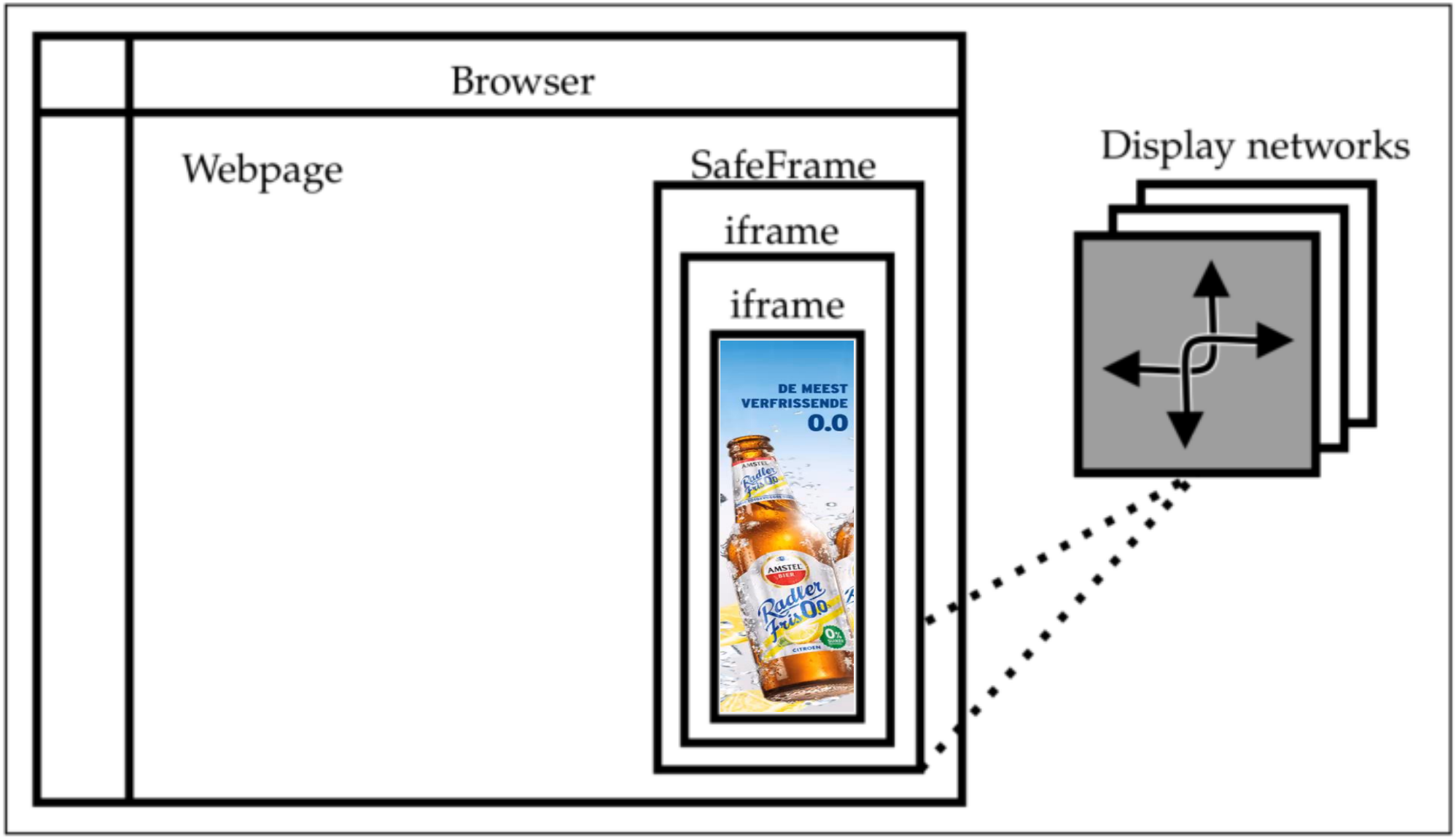


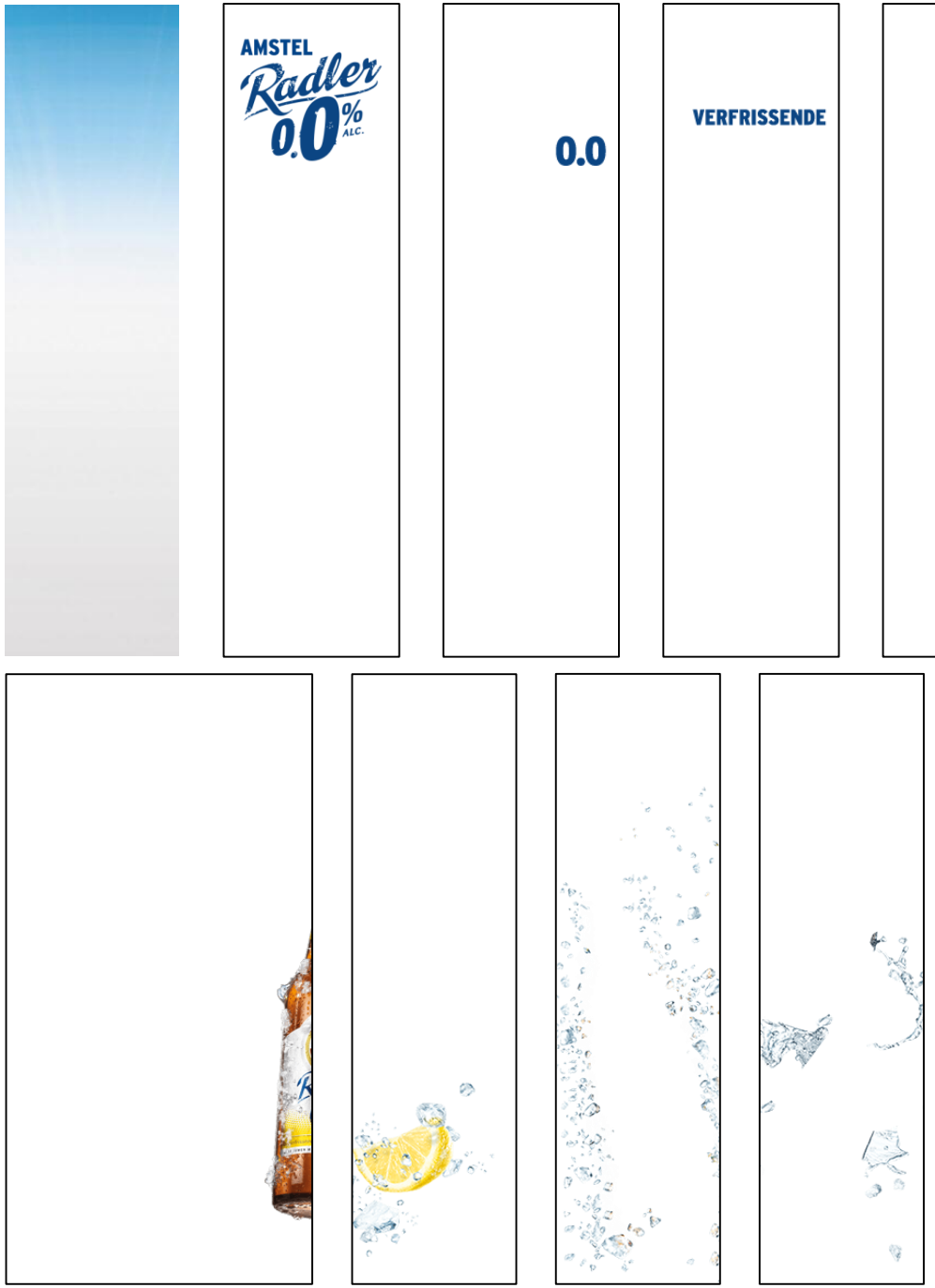
Two similar techniques:

- **Tracking pixel** (1x1 pixel, 0x0 pixel)
- **Fingerprint** (e.g. with the popular library '*fingerprintjs2*', URL: github.com/Valve/fingerprintjs2/)

```
var e = function(t) {
    if (!(this instanceof e)) return new e(t);
    this.options = this.extend(t, {
        swfContainerId: "fingerprintjs2",
        swfPath: "flash/compiled/FontList.swf",
        detectScreenOrientation: !0,
        sortPluginsFor: [/palemoon/i],
        userDefinedFonts: []
    }), this.nativeForEach = Array.prototype.forEach, this.nativeMap = Array.prototype.map
};
```

Behind the scenes of an online advertisement





alle	cookie	CSS	afb.	media	script	XHR	frame	overig
Huidige domein								
2mdn.net								
s0.2mdn.net		1	14		5			
ajax.googleapis.com					1			

Google Doubleclick Frame:
https://s0.2mdn.net/9026094/1559035272353/amnet_160x600/index.html

alle	cookie	CSS	afb.	media	script	XHR	frame	overig
Huidige domein								
weborama.fr	1							
cstatic.weborama.fr	1	9	36		2			
adrcdn.com								
media.adrcdn.com					2			
cloudflare.com								
cdnjs.cloudflare.com					2			
lemonpi.io								
d.lemonpi.io			1		1			

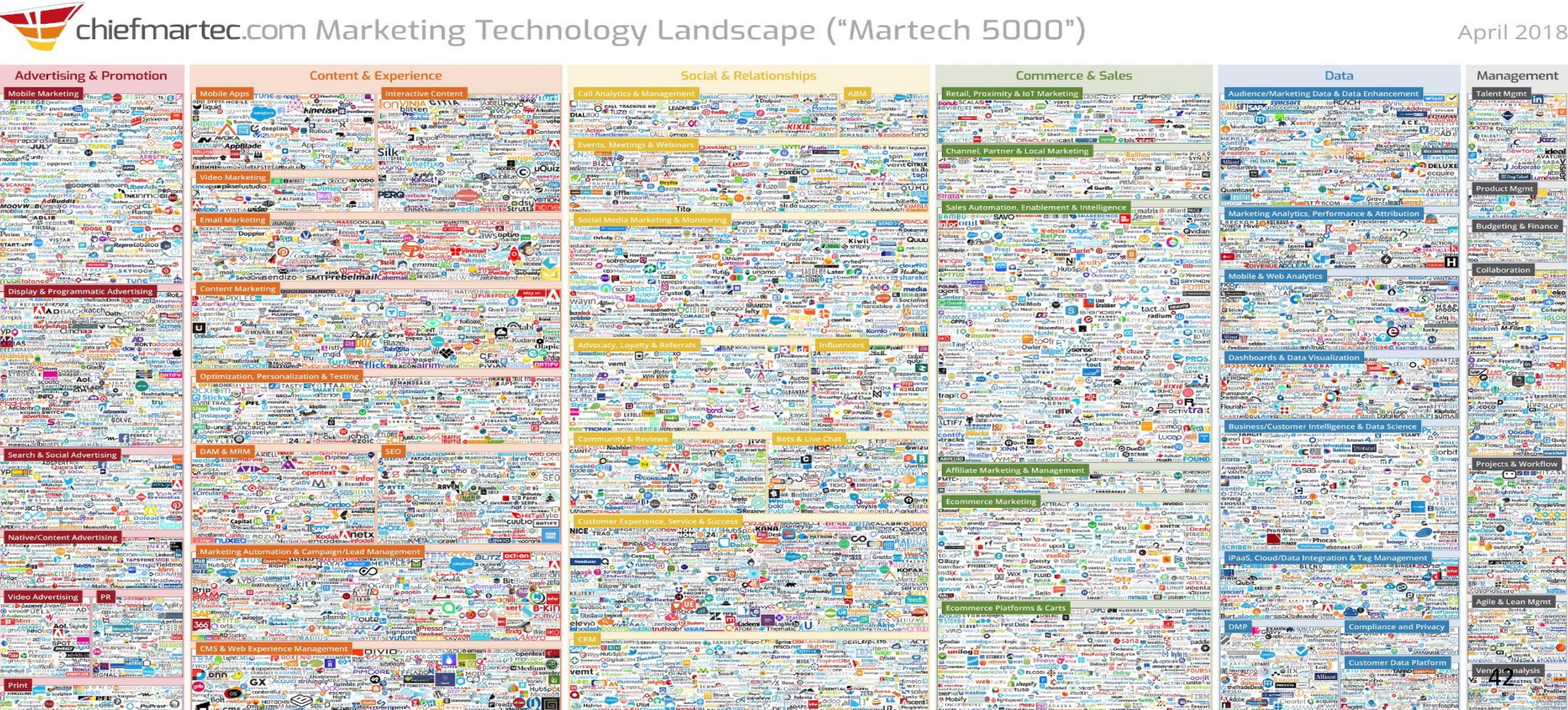
Weborama iframe:
https://cstatic.weborama.fr/advertiser/6760/2/3/8/weborama_apto_billboard_index.html?scrrefstr=scr_76599592073weborama_apto_billboard_index_html1561367308190&scrdebug=0&scrwidth=970&scrheight=250&scrwebodomain=0&scrdevtype=**desktop**&vars=wuid%3D%26**retargeting**%3D%26

Deeper behind the scenes!

Source: https://www.improvedigital.com/the-new-display-advertising-market-map-for-the-netherlands/

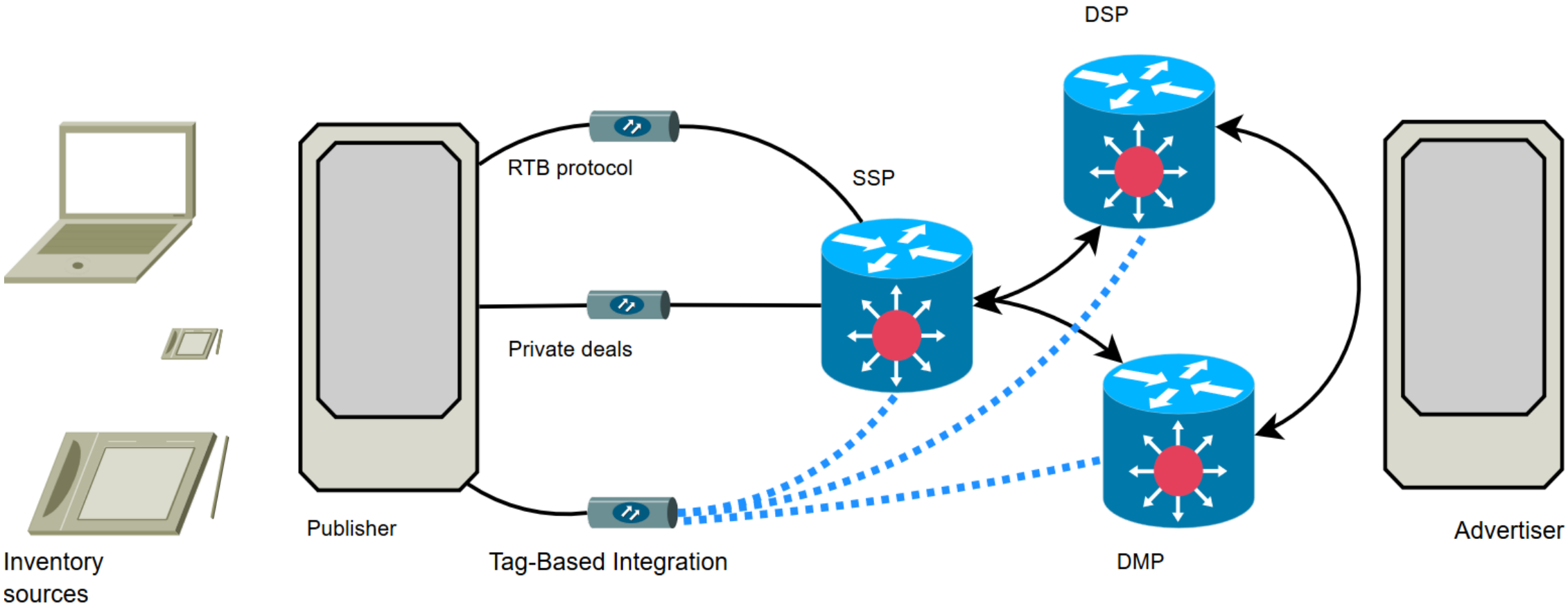
Source: https://chiefmartec.com/2019/04/marketing-technology-landscape-supergraphic-2019/

Even deeper behind the scenes!



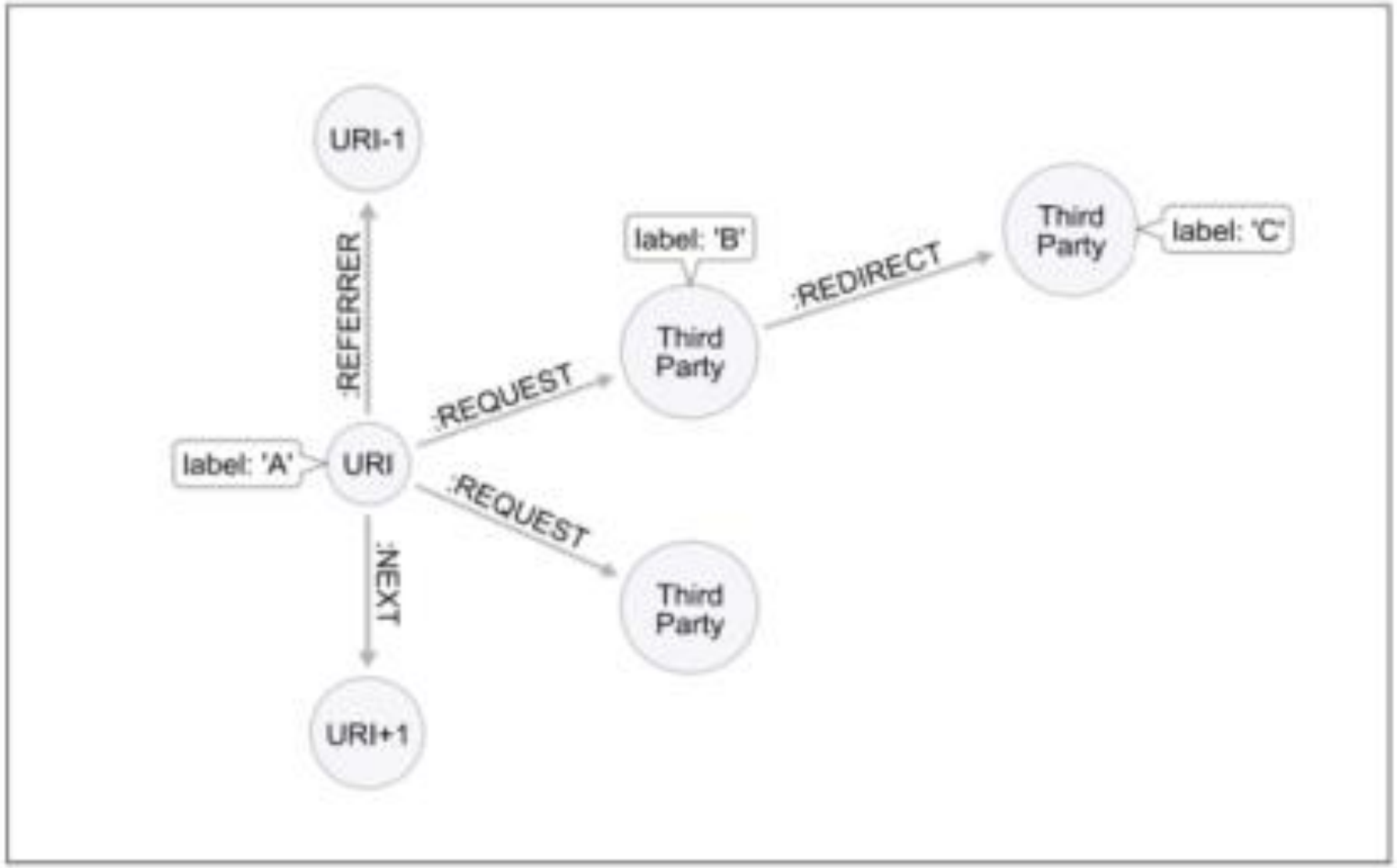
Video credits: https://nos.nl/video/527311-handel-in-een-fractie-van-een-seconde.html

A theoretical model for Real-Time Bidding

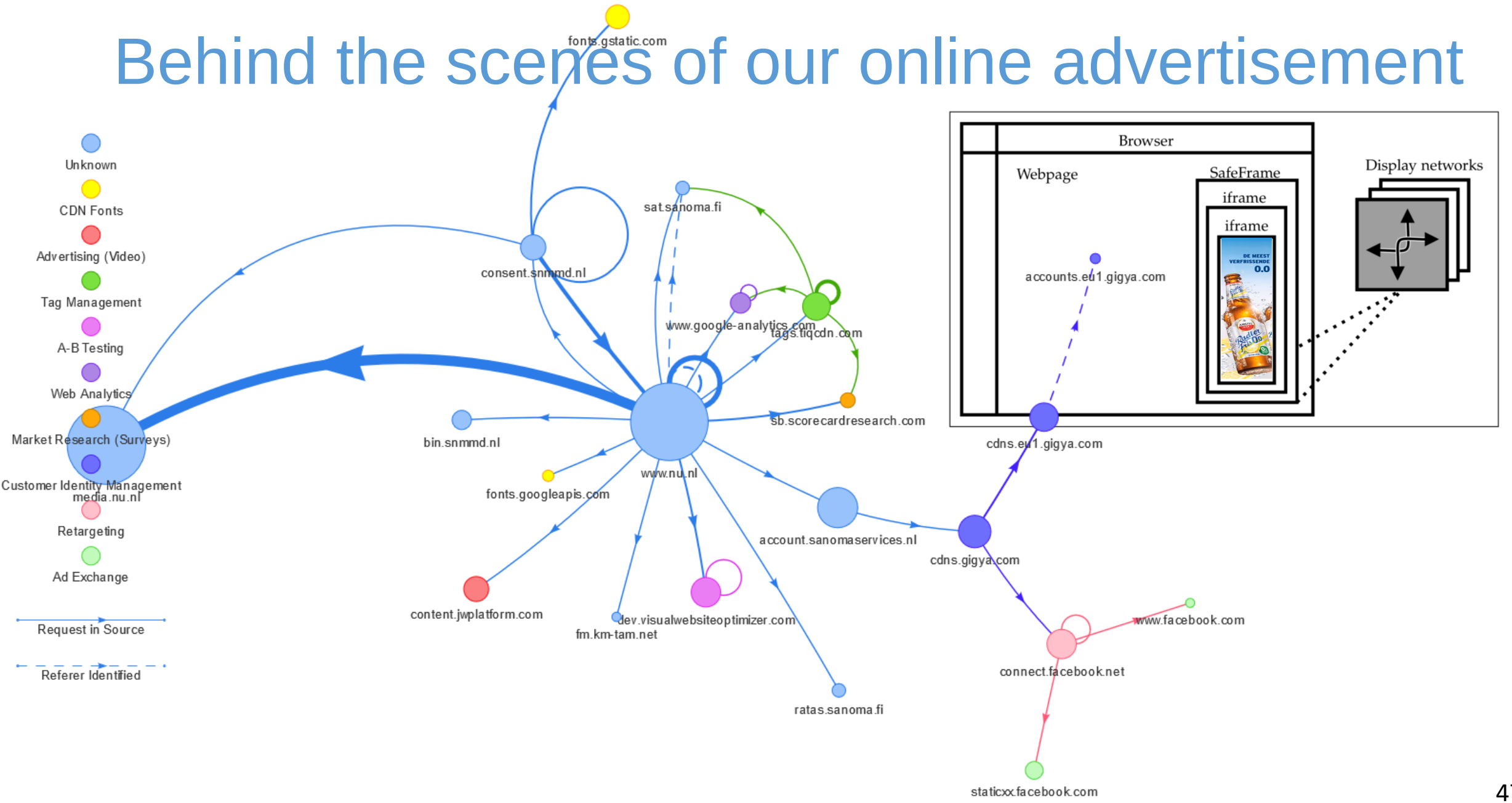


Bron: https://app.nos.nl/op3/cookies/

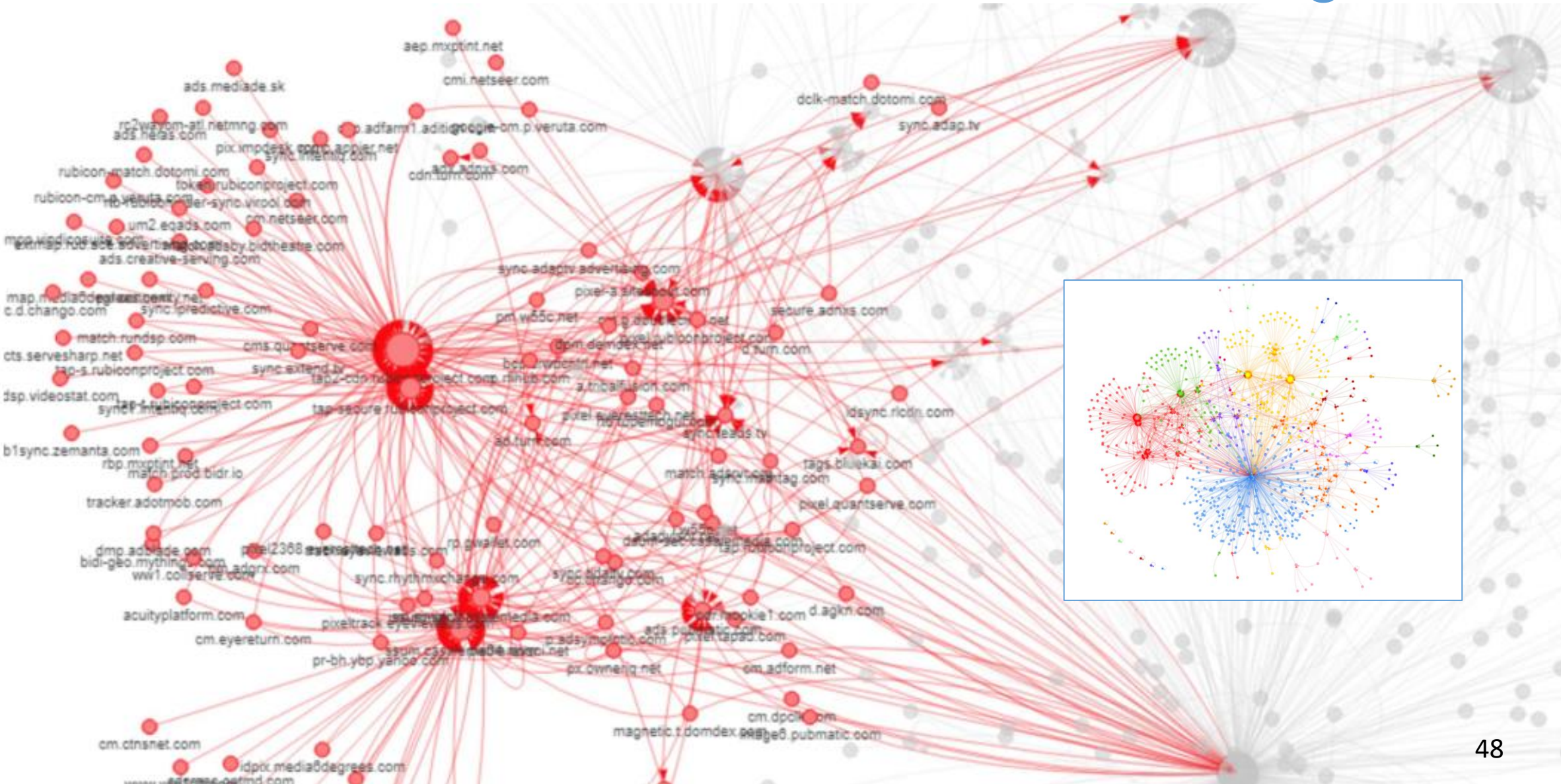
An empirical model for Real-Time Bidding



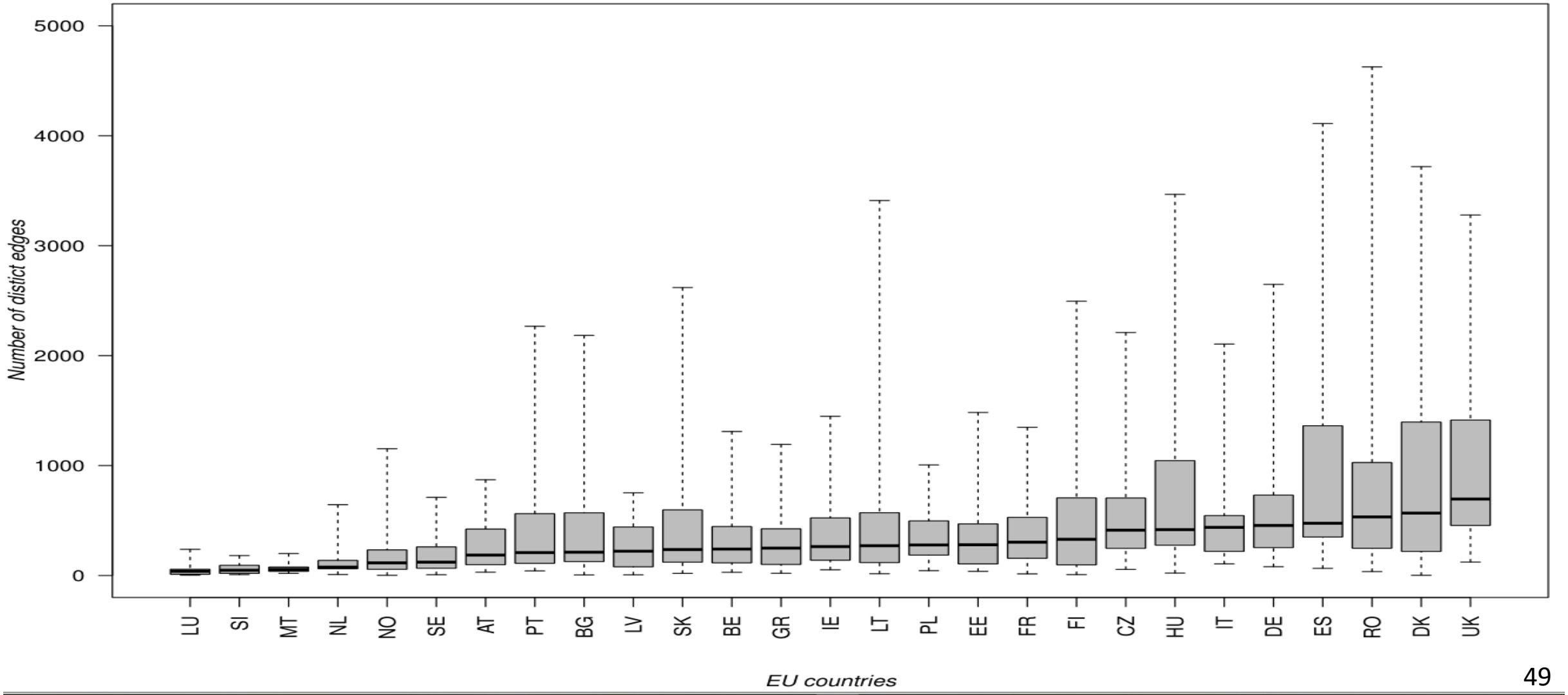
Behind the scenes of our online advertisement



Behind the scenes of Real-Time Bidding



Prior consent is a difficult point with cookies

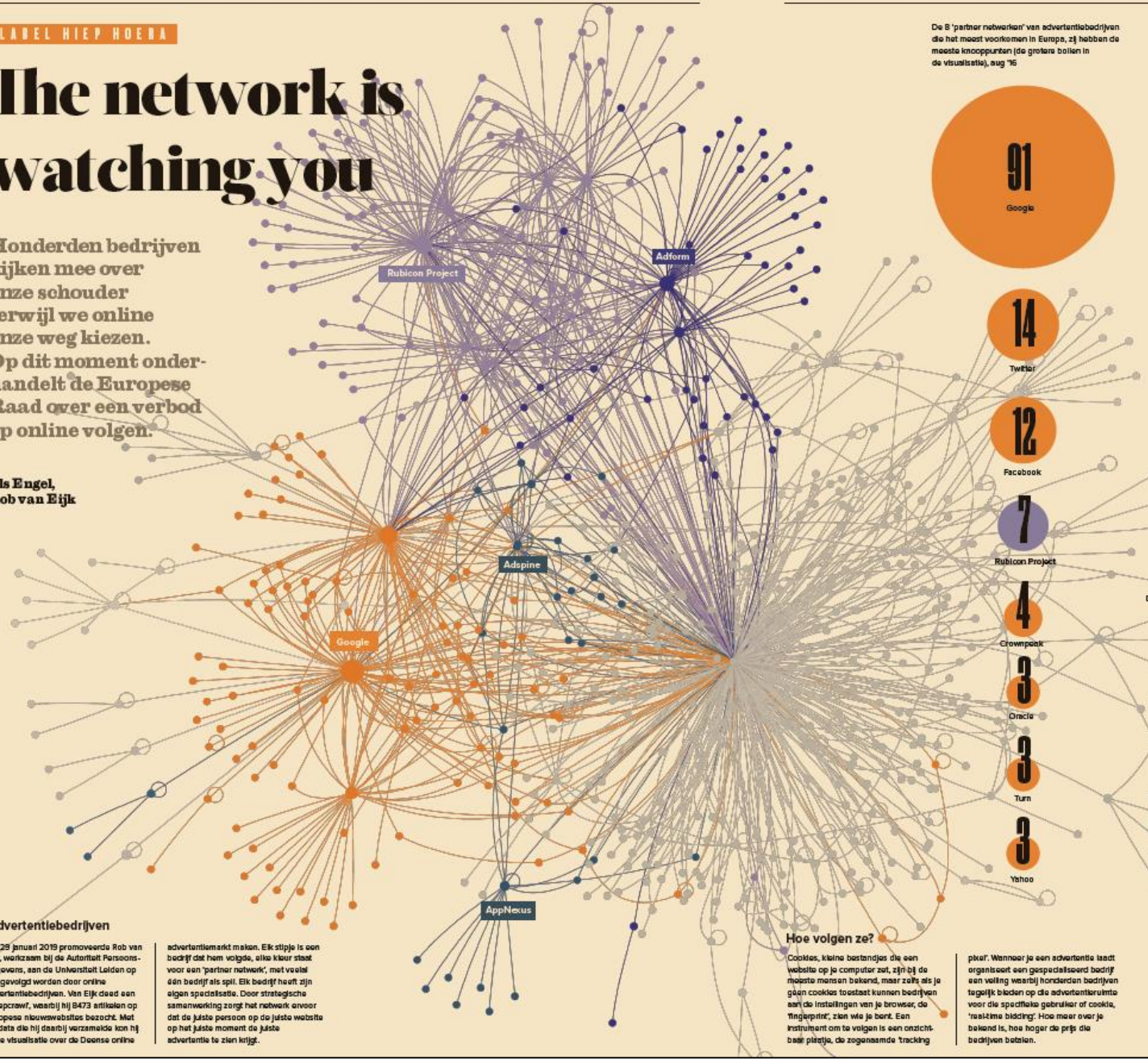


LABEL HIEP HOERA

The network is watching you

Honderden bedrijven kijken mee over onze schouder terwijl we online onze weg kiezen. Op dit moment onderhandelt de Europese Raad over een verbod op online volgen.

Els Engel, Rob van Eijk



De 8 'partner netwerken' van advertentiebedrijven die het meest voorkomen in Europa, zij hebben de meeste knooppunten (de grotere bollen in de visualisatie), aug '16

Rubicon Project

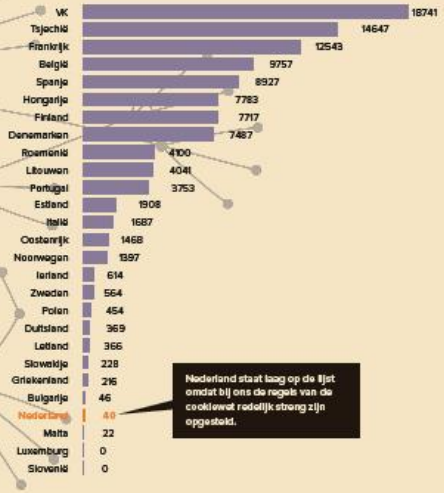
Dr. Van Eijk heeft sinds 2011 meerdere 'crawls' uitgevoerd. De top 3 advertentiebedrijven die hij het meeste tegenkwam in Europa bestaat uit bedrijven die ook diensten aan consumenten bieden. Het eerste achte advertentiebedrijf is Rubicon Project. Dit bedrijf uit de VS is in 2007 opgericht en heeft een omzet van $155,5 mln. Door overnames, fusies en strategische samenwerking is het Rubicon netwerk enorm gegroeid.

Aantal partners in het Rubicon Project netwerk

Grote verschillen binnen de EU

Er zijn weliswaar Europese regels voor cookies, maar elk land interpreteert deze regels anders, daarom zijn er andere mogelijkheden voor advertentiebedrijven. De EU werkt daarom nu aan de 'ePrivacy Verordening', waardoor het straks in elk land verboden is om online te volgen, tenzij expliciet toestemming is gegeven. De Voorzitter van de Europese Raad lijkt veel te willen maken met de nieuwe regels gezien de verklaringen van het Europees Parlement in mei van dit jaar.

Aantal keren dat een partner van Rubicon Project voorkwam in aug '16, per EU-land



Advertentiebedrijven

Op 29 januari 2019 promoveerde Rob van Eijk, werkzaam bij de Autoriteit Persoonsgegevens, aan de Universiteit Leiden op het gevolgd worden door online advertentiebedrijven. Van Eijk deed aan 'deepcrawl', waarbij hij 8473 artikelen op Europese nieuwswebsites bezocht. Met de data die hij daarbij verzamelde kon hij deze visualisatie over de Deense online advertentiemarkt maken. Elk stipje is een bedrijf dat hem volgde, elke kleur staat voor een 'partner network', met veelal één bedrijf als spil. Elk bedrijf heeft zijn eigen specialisatie. Door strategische samenwerking zorgt het network ervoor dat de juiste persoon op de juiste website op het juiste moment de juiste advertentie te zien krijgt.

Hoe volgen ze?

Cookies, kleine bestandjes die een website op je computer zet, zijn bij de meeste mensen bekend, maar zelfs als je geen cookies toestaat kunnen bedrijven aan de instellingen van je browser, de 'fingerprint', zien wie je bent. Een instrument om te volgen is een onzichtbaar plaatje, de zogenaamde 'tracking pixel'. Wanneer je een advertentie laadt organiseert een gespecialiseerd bedrijf een veiling waarbij honderden bedrijven tegelijk bieden op die advertentieruimte voor die specifieke gebruiker of cookie, 'real-time bidding'. Hoe meer over je bekend is, hoe hoger de prijs die bedrijven betalen.

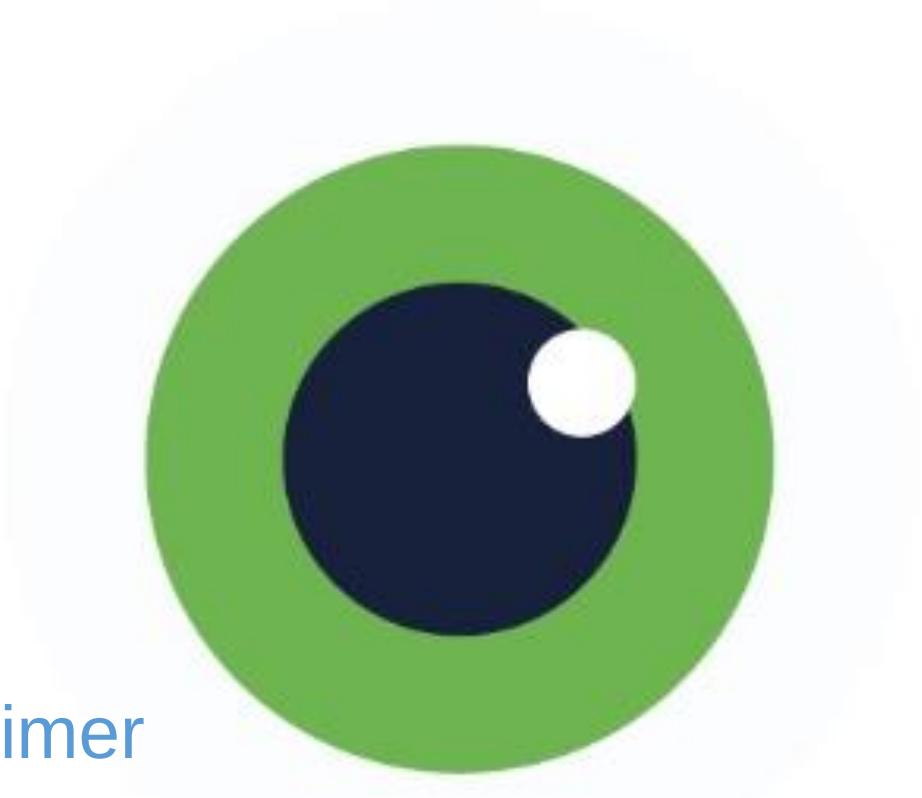
Mag dit?

Het is nog maar de vraag of alles wat deze advertentiebedrijven doen volgens de wet is. De EU-regels zeggen bijvoorbeeld dat je je data mag opvragen bij elk bedrijf, maar als je niet weet dat een bedrijf je volgt kan dat niet. Verder is het bij veel websites niet duidelijk hoe je het volgen kunt voorkomen of stopzetten.

Wat kun je zelf doen?

Goal regelmatig trackingcookies weg (Shift-Ctrl-Delete) tijdens het surfen. Of stel je browser goed in. De Firefox-browser kan bijvoorbeeld zo ingesteld worden dat deze trackers en tracking cookies blokkeert. Er zijn ook programma's die helpen bij het blokkeren van trackers, zoals bijvoorbeeld Ghostery.

Source: Het Financieele Dagblad, zaterdag 2 maart 2019



Disclaimer

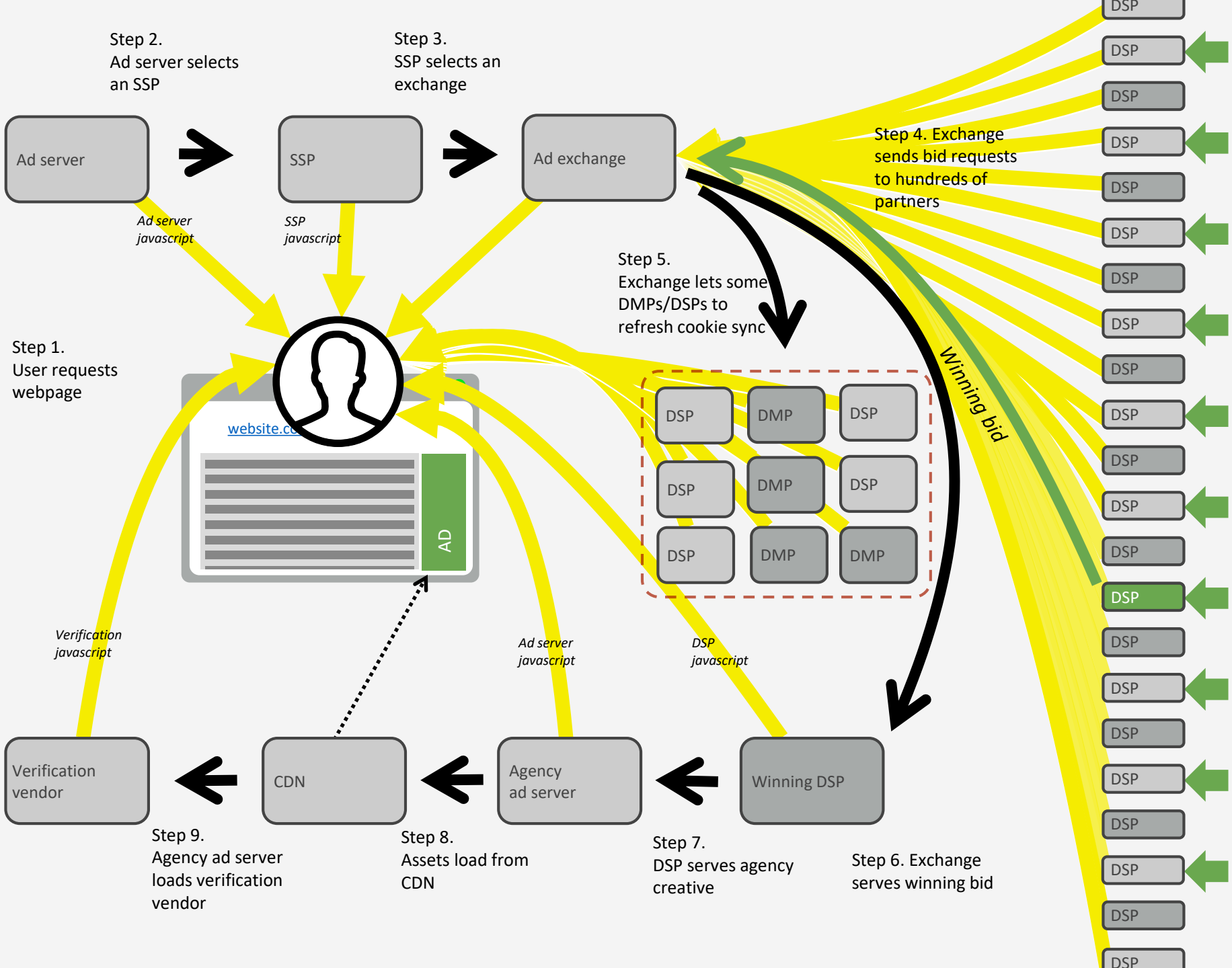
Animation with permission of dr. Johnny Ryan (Brave Browser).

DATA LEAKAGE IN ONLINE ADVERTISING

This is the current process of real-time bidding that is used in online behavioural advertising.

Legend

- Channel of data leakage
- Money



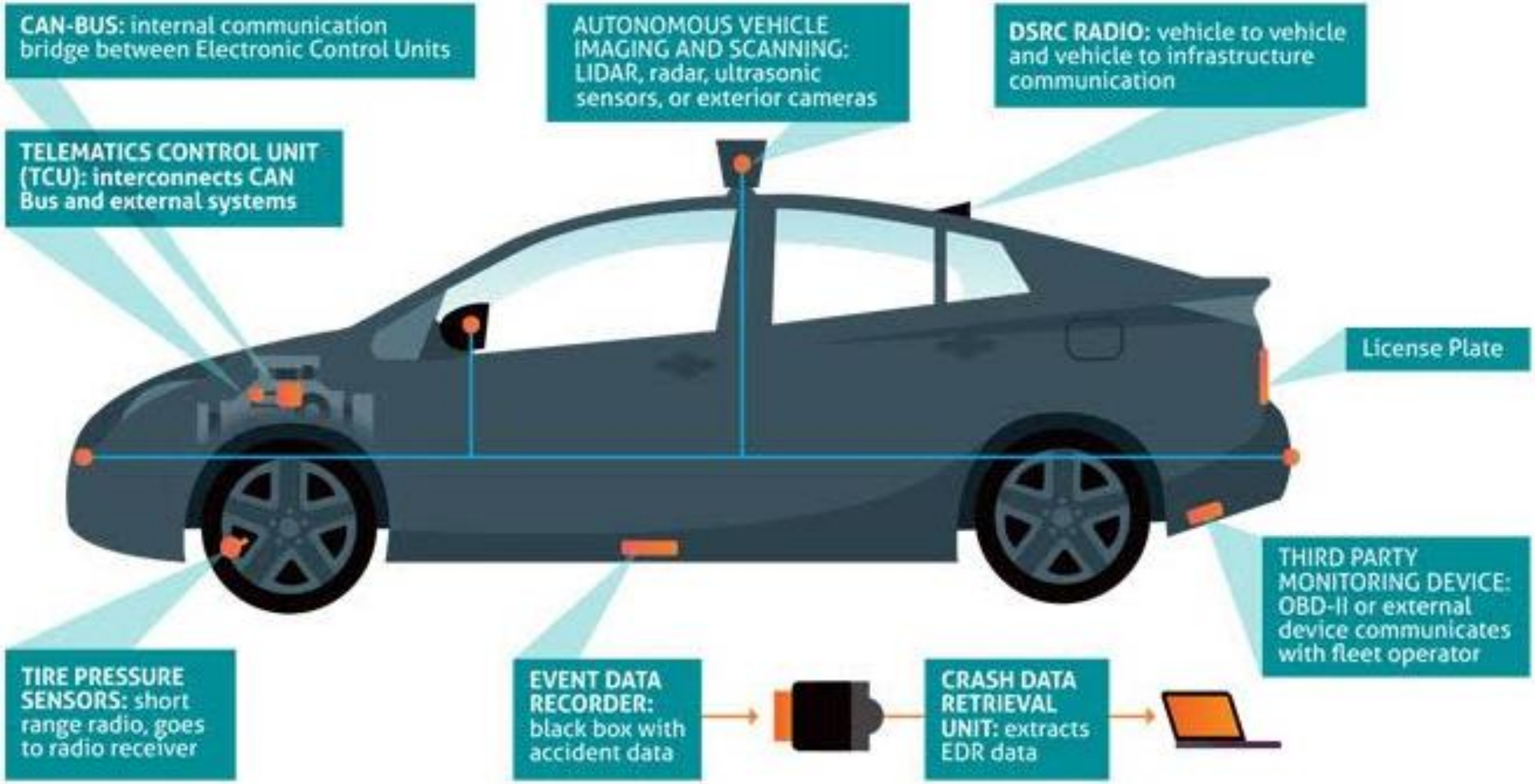
brave

MARKETERS

How do you view offline tracking?

- *"Data about driving behavior (braking, acceleration, switching energy, etc.)"*
- *"Data about travel behavior (I use Google Maps)"*
- *"Data about viewing behavior (I have a smart television)"*
- *“Data about brushing behavior (I have a smart toothbrush or vacuum cleaner”*
- *"Other IoT devices, such as Cayla doll"*

Offline tracking



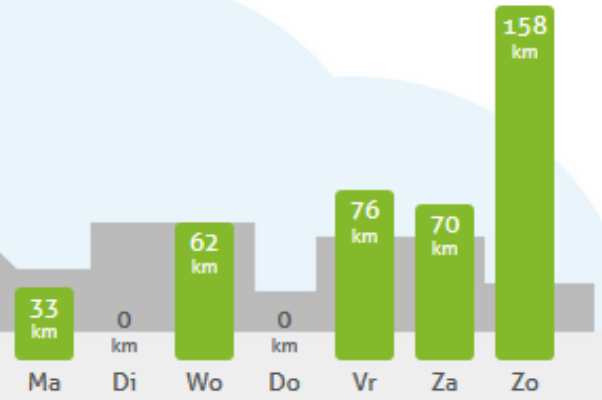
Source: https://fpf.org/issues/connected-

Overzicht van

Kilometeradministratie

17-06-2019 - 24-06-2019

Gemiddeld per dag



8 Steden bezocht

Rotterdam	7x
Zoetermeer	4x
's-Gravenhage	3x

Utrecht — Den Haag

Langste afstand 60km

Langste duur 01:00

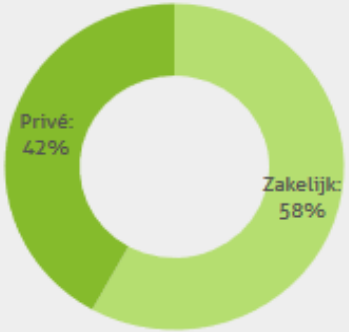
Rotterdam — Den Haag

Gereden

22 Routes

400 km

10 uren 51 min



4x een nuttige plaats bezocht

2x Huis van ... -22 dagen, -...

2x Mantelzor... -22 dagen, -...

1 Bestuurders

22x

Geparkeerd

7 dagen 11 uren 22 min

Kilometerstand

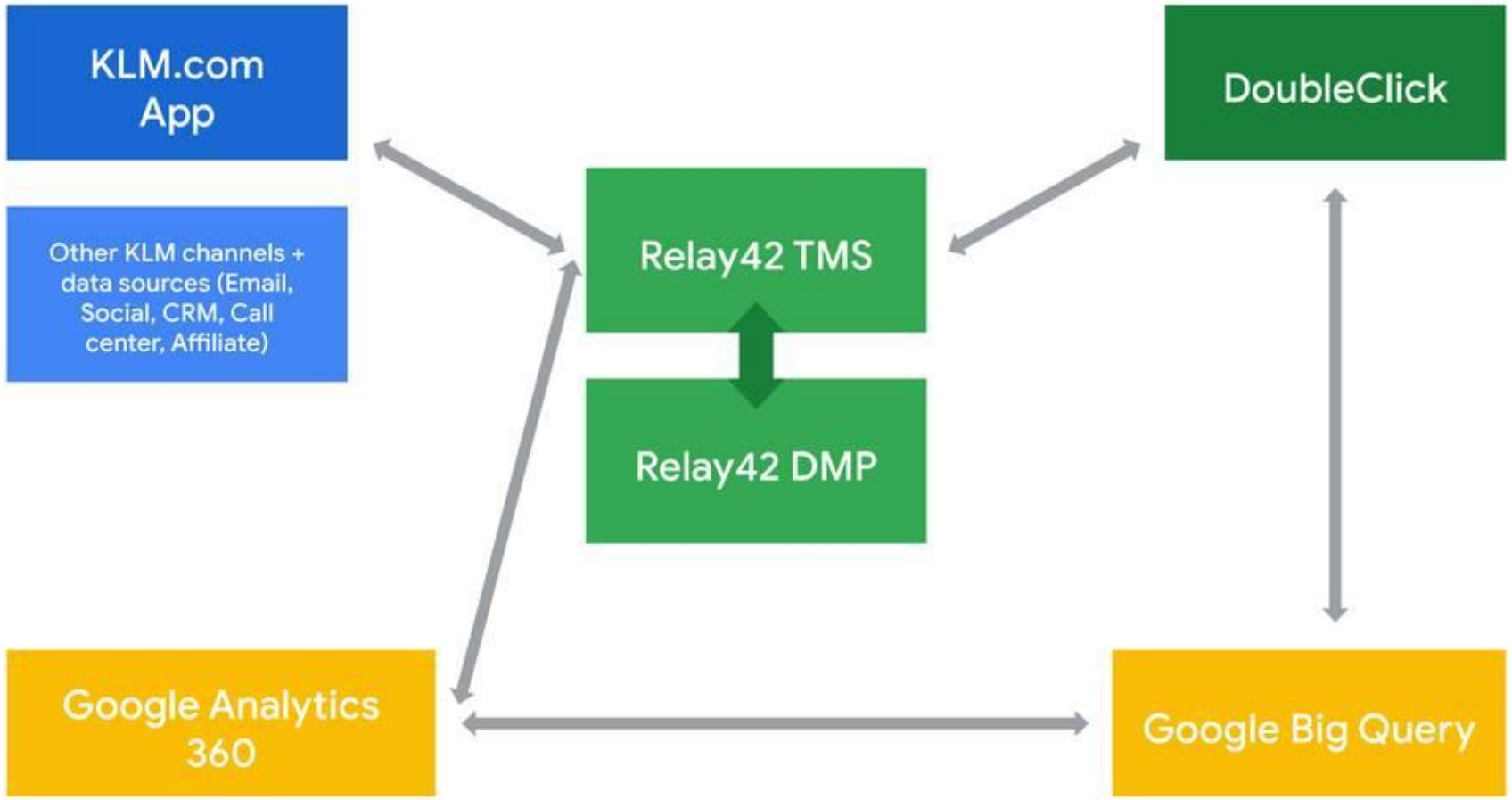
Profiling based on online and offline data (1)

- A Data Management Platform (DMP) can enrich online data with offline (personal) data.
- A DMP shares (profile) data with companies in the advertising chain.

Profiling based on online and offline data (2)

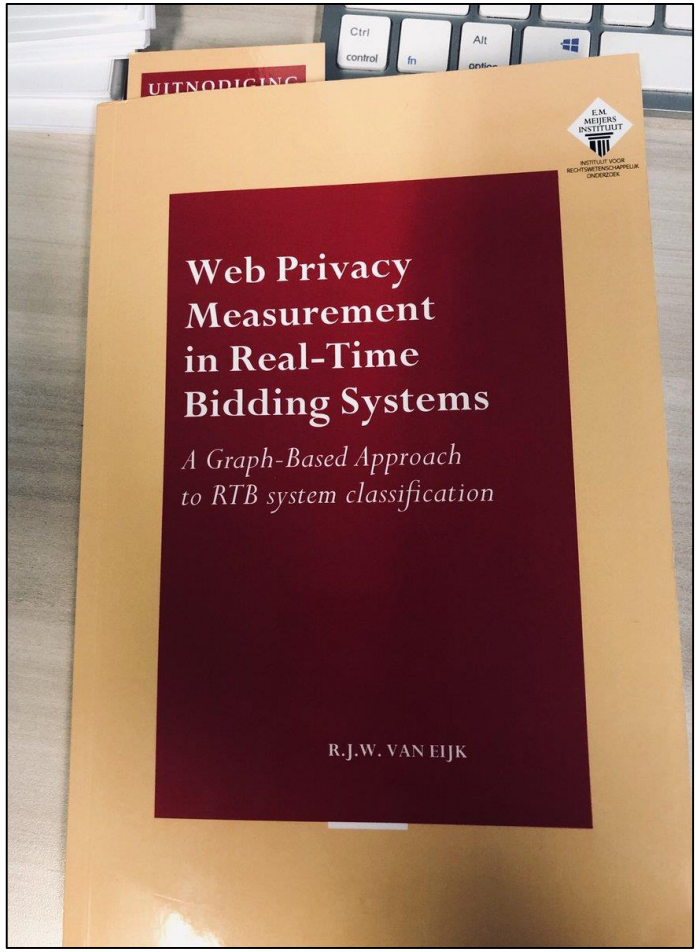
- A DMP thus enables marketers to zoom in on customers in real time.
- A DMP specializes in customer data.
- An example of a well-known DMP is Bluekai: "With more than 30 data suppliers, marketers have access to nearly 700 million customer profiles and 40,000 data attributes."

Profiling based on online and offline data (3)



Source: https://storage.googleapis.com/twg-content/images/klm_chartv32.width-1000.jpg

Thank you!



https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3319284

https://www.youtube.com/watch?v=YWJ7ZXEmzHw

Questions?

5 February 2020 | I-Interim Rijk Data Science Crash Course | The Hague

FUTURE OF PRIVACY FORUM