



# Exploratory study on Information Technology (IT) and Artificial Intelligence (AI) Tools for Monitoring Online Markets for Consumer Policy Purposes

European Commission Tender

Final Report

(JUST/2018/CONS/PR/CO01/0123)

*The information and views set out in this report are those of the authors and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.*

***European Commission Tender  
(JUST/2018/CONS/PR/CO01/0123)***



# **Exploratory study on Information Technology (IT) and Artificial Intelligence (AI) Tools for Monitoring Online Markets for Consumer Policy Purposes**

European Commission Tender  
Final Report

(JUST/2018/CONS/PR/CO01/0123)

*The information and views set out in this report are those of the authors and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use which may be made of the information contained therein.*

*This report is an excerpt from a full report prepared for the European Commission and exploring IT/AI tools for monitoring online markets for consumer policy purposes. Selected parts of the full report such as a list of tools used by consumer authorities as well as the exploratory assessment of the tools were omitted for reasons of confidentiality.*

## **Executive summary**

The purpose of this study is 'to identify and make an inventory of information technology (IT)/artificial intelligence tools (AI) that are or can be made useful in online market surveillance for consumer policy development purposes and for the enforcement of consumer protection legislation'.

In order to find the relevant tools we distributed questionnaires to researchers as well as consumer authorities and organizations to find out what are the tools that they know and we searched online databases. We also included the tools that we learned about via (electronic) 'word of mouth' that is either by talking to other researchers and experts or through social media. This resulted in a list of 517 potentially relevant tools. To make sure that the list includes only tools that are indeed relevant to this study, we conducted an individual check. Each tool was coded independently by three members of our research group who indicated whether the tool is relevant or not. A tool was included in the final list of relevant tools if it was coded as 'relevant' by all three researchers. 73 tools made it to the final list.

The tools reported by consumer authorities are of a general nature, i.e., they can be used for multiple purposes also outside the scope of the tasks of consumer authorities. Examples are VPN tools, registries and tools to collect evidence from the websites and/or to track changes of website content. Exceptionally, consumer law authorities developed tools themselves that are specifically targeted to their own needs.

Most of the tools developed by researchers deal with issues such as the detection of dark patterns, the non-disclosure of affiliate marketing by social media influencers, Facebook ad profiling or price discrimination. What is very specific about these tools is that they generally do not have an interface that would make them available for use by a layperson.

In contrast, the tools found in software registries, as well as some of the tools that we have heard about via word of mouth, are tools aiming at consumer empowerment, which means that they are meant to be used by consumers. Examples of tools of this kind are Ghostery or Scamadviser.

Generally, the tools for monitoring infringements of consumer law that are ready for deployment are relatively few. Distill probably comes closest to the goal of a tool monitoring websites for a number of infringements of consumer law. Similarly, the Dark Patterns study could be used at regular intervals to monitor and record the activity of a considerable amount of online stores. However, given that this tool does not have any interface at the moment, it cannot be deployed by a consumer protection authority without personnel with programming skills. Claudette (unfair terms checker) can be adapted for monitoring.

In addition, a number of tools can be helpful to search for infringements of consumer law. Reviewmeta's list of least trusted reviews or Scamadvisor's list of biggest scams can, for example, provide an indication that further investigation is required.

Furthermore, some tools that allow for individual investigations into specific issues (e.g. finding the IP address of a given trader) or to collect evidence can also be useful for consumer authorities.

One of the results we find promising is that the majority of tools can be described as highly scalable. This implies that there is a potential for extending the use of these tools beyond their current scope.

The question as to which tools to use or to choose cannot be answered in isolation. The tools identified in this study depict the current state of affairs in terms of tool development. Focusing on how the tools ought to be used, shared and implemented by consumer authorities is a very important short-term focus. However, it is also necessary to reflect on how a broader strategy could look like in the long term. Public interest technology can take the form of in-house expertise (authorities developing data units), grassroots activity (civil society developing tools for consumer empowerment), public industry tools (industry associations develop tools to avoid market failures), or academic research (tools developed as a result of academic research). When considering the risks and benefits of these approaches, three aspects need to be emphasized: data ownership and privacy, ethics, and cybersecurity. Issues relating to data ownership and privacy arise primarily out of the combination of malicious and / or commercial interests and the business models of the new data economy, which remain largely unexplored (e.g. consumer authorities may have no control over data collected by commercial providers of AI/IT tools, such as VPN services). With respect to ethics, monitoring activities need to be legitimized by the legal mandate of consumer authorities at the national and European level. In terms of cybersecurity, in performing digital monitoring and investigations through the use of tools built or purchased for the purposes of enforcing consumer protection, it is of utmost importance to consider informational security from two main perspectives. First, public authorities need to be aware that cyberattacks are becoming more prevalent, and need to take measures to protect themselves. Second, consumer law enforcement often overlaps with criminal law, which calls for more effective coordination between consumer and law enforcement authorities. Enhanced cooperation ought to be considered also for information sharing with other national and EU authorities responsible for competition law and data protection, as well as for infrastructural investment in IT expertise, with a specific emphasis on interoperability.

## **Sommaire**

L'objectif de cette étude est "d'identifier et de dresser un inventaire des technologies de l'information (TI) et des outils d'intelligence artificielle (IA) qui sont ou peuvent être utiles dans la surveillance du marché en ligne à des fins d'élaboration de la politique des consommateurs et d'application de la législation relative à la protection des consommateurs".

Afin de trouver les outils pertinents, nous avons distribué des questionnaires aux chercheurs ainsi qu'aux autorités et organisations de consommateurs pour savoir quels sont les outils qu'ils connaissent et nous avons effectué des recherches dans des bases de données en ligne. Nous avons également inclus les outils dont nous avons eu connaissance par le "bouche à oreille" (électronique), c'est-à-dire en parlant à d'autres chercheurs et experts ou par le biais des médias sociaux. Nous avons ainsi obtenu une liste de 517 outils potentiellement pertinents. Pour nous assurer que la liste ne comprend que des outils qui sont effectivement pertinents pour cette étude, nous avons procédé à une vérification individuelle. Chaque outil a été codé indépendamment par trois membres de notre groupe de recherche qui ont indiqué si l'outil était pertinent ou non. Un outil a été inclus dans la liste finale des outils pertinents s'il a été codé comme "pertinent" par les trois chercheurs. 73 outils ont été inclus dans la liste finale.

Les outils signalés par les autorités chargées de la protection des consommateurs sont de nature générale, c'est-à-dire qu'ils peuvent être utilisés à des fins multiples, même en dehors du cadre des tâches des autorités chargées de la protection des consommateurs. Il s'agit par exemple des outils VPN, des registres et des outils permettant de recueillir des preuves sur les sites web et/ou de suivre les modifications du contenu des sites web. Exceptionnellement, les autorités chargées de la protection des consommateurs ont développé elles-mêmes des outils spécifiquement adaptés à leurs propres besoins.

La plupart des outils développés par les chercheurs traitent de questions telles que la détection de schémas sombres, la non-divulgaration du marketing d'affiliation par les influenceurs des médias sociaux, le profilage des publicités sur Facebook ou la discrimination par les prix. Ces outils sont très spécifiques car ils ne disposent généralement pas d'une interface qui les rendrait accessibles à un profane.

En revanche, les outils que l'on trouve dans les registres de logiciels, ainsi que certains des outils dont nous avons entendu parler par le bouche à oreille, sont des outils visant à responsabiliser les consommateurs, ce qui signifie qu'ils sont destinés à être utilisés par les consommateurs. Ghostery ou Scamadviser sont des exemples d'outils de ce type.

En général, les outils de suivi des infractions au droit de la consommation qui sont prêts à être déployés sont relativement peu nombreux. Distill est probablement l'outil qui se rapproche le plus de l'objectif d'un outil de surveillance des sites web pour un

nombre d'infractions au droit de la consommation. De même, l'étude Dark Patterns pourrait être utilisée à intervalles réguliers pour surveiller et enregistrer l'activité d'un nombre considérable de magasins en ligne. Toutefois, étant donné que cet outil ne dispose pas d'interface pour le moment, il ne peut être déployé par une autorité de protection des consommateurs sans un personnel ayant des compétences en programmation. Claudette (contrôleur de clauses abusives) peut être adapté pour la surveillance.

D'autres d'outils peuvent être utiles pour détecter des infractions au droit de la consommation. La liste des examens les moins fiables de Reviewmeta ou la liste des plus grandes escroqueries de Scamadvisor peuvent, par exemple, indiquer qu'une enquête plus approfondie est nécessaire.

En plus, certains outils qui permettent de mener des enquêtes individuelles sur des questions spécifiques (par exemple, trouver l'adresse IP d'un commerçant donné) ou de recueillir des preuves peuvent également être utiles aux autorités chargées de la protection des consommateurs.

L'un des résultats que nous trouvons prometteurs est que la majorité des outils peuvent être décrits comme hautement évolutifs. Cela implique qu'il est possible d'étendre l'utilisation de ces outils au-delà de leur portée actuelle.

La question de savoir quels outils utiliser ou choisir ne peut être résolue de manière isolée. Les outils identifiés dans cette étude décrivent l'état actuel des choses en termes de développement d'outils. Il est très important à court terme de se concentrer sur la manière dont les outils devraient être utilisés, partagés et mis en œuvre par les autorités chargées de la protection des consommateurs. Cependant, il est également nécessaire de réfléchir à la manière dont une stratégie plus large pourrait être mise en place à long terme. Les technologies d'intérêt public peuvent prendre la forme d'une expertise interne (les autorités développant des unités de données), d'une activité de base (la société civile développant des outils pour la responsabilisation des consommateurs), d'outils de l'industrie publique (les associations industrielles développant des outils pour éviter les défaillances du marché) ou de la recherche universitaire (outils développés à la suite de la recherche universitaire).

Lors de l'examen des risques et des avantages de ces approches, trois aspects doivent être soulignés: la propriété des données et la vie privée, l'éthique et la cybersécurité. Les questions relatives à la propriété des données et à la vie privée découlent principalement de la combinaison d'intérêts malveillants et/ou commerciaux et des modèles commerciaux de la nouvelle économie des données, qui restent largement inexplorés (par exemple, les autorités chargées de la consommation peuvent n'avoir aucun contrôle sur les données collectées par les fournisseurs commerciaux d'outils AI/TI, tels que les services VPN). En ce qui concerne l'éthique, les activités de contrôle doivent être légitimées par le mandat légal des autorités de protection des consommateurs au niveau national et européen. En ce qui concerne la cybersécurité, en effectuant une surveillance et des enquêtes numériques à l'aide



d'outils construits ou achetés dans le but de faire respecter la protection des consommateurs, il est de la plus haute importance d'envisager la sécurité de l'information sous deux angles principaux. Premièrement, les autorités publiques doivent être conscientes que les cyberattaques sont de plus en plus fréquentes et doivent prendre des mesures pour se protéger. Deuxièmement, l'application du droit de la consommation recoupe souvent le droit pénal, ce qui nécessite une coordination plus efficace entre les autorités chargées de la protection des consommateurs et celles chargées de l'application de la loi. Une coopération renforcée devrait également être envisagée pour le partage d'informations avec d'autres autorités nationales et européennes responsables du droit de la concurrence et de la protection des données, ainsi que pour les investissements infrastructurels dans l'expertise informatique, avec un accent particulier sur l'interopérabilité.

## **Abstract**

The study aims to identify and make an inventory of IT and AI that are or can be made useful in online market surveillance for consumer policy development purposes and for the enforcement of consumer protection legislation.

The study shows that relatively few tools exist that are ready to use for monitoring infringements of consumer law. Distill seems to come closest to this goal. The Dark Patterns study could also be used at regular intervals to monitor and record the activity of a considerable amount of online stores. However, since it has no user interface, sound programming skills are required to use it. In addition, tools such as Reviewmeta and Scamadvisor provide lists of suspicious cases that can guide authorities in deciding on cases that require further investigation. Furthermore, tools exist for finding information about specific IP addresses, VAT-numbers etc. and to collect evidence of infringements. The report also briefly touches upon tools for policy development such as Sense4Us.

Finally, the report makes suggestions for a general framework on how to use and share existing tools and on elements to take into account when considering the development of tools tailored to the needs of consumer authorities.

## Table of Contents

1	Introduction	4
1.1	Background	4
1.2	Aim of the study	5
1.3	Structure	6
2	Inventory of AI / IT tools	7
2.1	Introduction	7
2.2	Methodology	7
2.2.1	Questionnaires	7
2.2.2	Database searches	8
2.2.3	Word of mouth	9
2.3	Filtering out irrelevant tools and interim results	9
2.4	Interim conclusions	10
3	Proposed criteria for tool assessments	11
3.1	Methodology: Selecting tools for further assessment	11
4	Recommendations	18
4.1	Reframing the need for technological solutions	18
4.2	Invest in cooperation and information sharing with other consumer authorities and in the tools that allow this to take place in an efficient way	25
4.3	Invest in cooperation and information sharing with criminal law, competition law and data protection enforcement agencies and discuss access to some of their AI/IT tools	26
4.3.1	Consumer law and criminal law	26
4.3.2	Consumer law and competition law	28
4.3.3	Consumer law and data protection law	28
4.3.4	Cooperate with national and EU consumer associations	28
4.4	Look into the possible use of (general) tools for policy development for the development of consumer policy	29
4.5	Use the shortlisted monitoring tools in order to obtain a first indication of possible infringements	29
4.6	Use the shortlisted investigation and evidence collection tools taking into account their strengths and weaknesses	30

4.7	Invest in the use and development of tools aimed specifically at the needs of consumer authorities with regard to the enforcement of consumer law	31
4.8	Accuracy issues with AI tools	31
4.9	Create European repositories for tools used by authorities	32
4.9.1	Public repositories	32
4.9.2	Internal repositories: interoperable monitoring tools and consumer complaints database infrastructure	35
4.9.3	Connect the monitoring tools dashboard with the consumer complaints database of each relevant EU consumer protection authority	37
5	List of literature	38
6	List of Annexes	40
Annex 2	Questionnaire for consumer agencies/organizations	41
Annex 3	Questionnaire for researchers	47
Annex 4	List of consumer agencies/organizations contacted	49
Annex 5	List of search strings and categories	50
Annex 6	Number of potentially relevant tools depending on the source	55
Annex 7	A list of 73 relevant tools	56

## **List of abbreviations**

AI	Artificial intelligence
API	Application Programming Interface
CEMP	Consumer Empowerment
CPC	Consumer protection cooperation
DNS	Domain name system
DSM	Digital Single Market
EC	European community
e.g.	Ex generis, for instance
EU	European Union
FACCT	Fairness, Transparency and Accountability
Ibid.	ibidem
i.e.	id est
IP	internet protocol
IT	Information technology
MS	Member States
SaaS	Software as a Product
SaaS	Software as a Service
ToR	Terms of Reference
UCP	Unfair commercial practices
UCT	Unfair contract terms

# 1 Introduction

## 1.1 Background

The completion of the Digital Single Market (DSM) is one of the political priorities of the European Union.<sup>1</sup> The Digital Single Market is defined as 'one in which the free movement of persons, services and capital is ensured and where the individuals and businesses can seamlessly access and engage in online activities under conditions of fair competition, and a **high level of consumer and personal data protection**, irrespective of their nationality or place of residence'.<sup>2</sup>

The digital single market is not only a political objective. It is part of the daily life of European citizens. The number of online consumer transactions increases every year. In 2018, for example, 69% of EU internet users shopped online.<sup>3</sup>

A high level of consumer and personal data protection for European citizens requires not only adequate legislation, but also compliance with such legislation. Effective and efficient enforcement is therefore of utmost importance.

In order to facilitate the enforcement of EU consumer law, the EU adopted the CPC Regulation<sup>4</sup> giving national enforcement authorities wider enforcement powers.

At the end of January, the European Commission published the results of its conformity check of 560 e-commerce websites. About 60% of them contained infringements.<sup>5</sup> These findings were the result of an EU wide sweep of e-Commerce websites. Sweeps are carried out every year on a specific theme and they are very labour and time intensive. There is an urgent need to facilitate and automate the finding of online infringements of consumer law. Ideally, a large number of websites offering goods or services to consumers, not only e-commerce, should be monitored continuously on an automated basis. This raises the question as to the availability of AI/IT tools that would enable such monitoring. Moreover, since businesses are increasingly using AI tools to manipulate consumers, it is important to start using AI to defend consumer rights.<sup>6</sup>

---

<sup>1</sup> For the period 2014-2019, see U. von der Leyen, Political guidelines for the next European Commission 2019-2024, [https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission\\_en.pdf](https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf)

<sup>2</sup> <https://ec.europa.eu/digital-single-market/en/policies/shaping-digital-single-market>.

<sup>3</sup> [https://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce\\_statistics\\_for\\_individuals](https://ec.europa.eu/eurostat/statistics-explained/index.php/E-commerce_statistics_for_individuals).

<sup>4</sup> Regulation (EU) 2017/2394 of the European Parliament and of the Council of 12 December 2017 on cooperation between national authorities responsible for the enforcement of consumer protection laws and repealing Regulation (EC) No 2006/2004, OJ L 345, 27.12.2017, p. 1–26.

<sup>5</sup> See [https://ec.europa.eu/commission/presscorner/detail/en/ip\\_20\\_156](https://ec.europa.eu/commission/presscorner/detail/en/ip_20_156), See also [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_19\\_1333](https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1333).

<sup>6</sup> Compare M. Lippi, G. Contissa, F. Lagioia, et al. (2019) Consumer protection requires artificial intelligence. Nat Mach Intell 1, 168–169 doi:10.1038/s42256-019-0042-3.

## 1.2 Aim of the study

Following the Terms of Reference, which states that 'Consumer policy and its enforcement by competent national authorities need new tools that contribute to broaden, simplify and automate its surveillance of online consumer markets, and the collection of electronic evidence from e-commerce websites, platforms and apps', this study identifies and makes an inventory of **information technology (IT)/artificial intelligence tools (AI)** that are or can be made useful in **online market surveillance for consumer policy development purposes and for the enforcement of consumer protection legislation**.<sup>7</sup>

From this, it follows that the tools we are looking for need to satisfy the following criteria:

1. **Nature:** AI/IT tools;
2. **Functionality:** they are or can be made useful for one of the following two purposes (that cannot always be clearly distinguished):
  - a. online market surveillance for consumer policy development purposes;
  - b. enforcement of consumer protection legislation.

The value of the results, and therefore this study, is the overview and assessment of tools that have been identified nor used by consumer protection authorities.

This study builds on a number of concepts that are central in this document. One of these concepts is '**online market surveillance**', which is defined here as **the activity of checking whether the activities of sellers or service providers that offer goods or services online comply with consumer protection legislation and do not give rise to consumer complaints, including whether the goods or services offered comply with the reasonable expectations of consumers**.

Furthermore, '**consumer policy development**' is understood as the policy with regard to:

- the development of new legislation or other types of regulation aimed at consumer protection, the modification and withdrawal of such legislation or regulation; and
- the development of strategies, priorities and techniques for the enforcement of existing consumer protection legislation.

Additionally, the '**enforcement of consumer protection legislation**' is understood to include public enforcement and private enforcement of consumer protection legislation. Although market surveillance by consumer protection authorities will generally lead to public enforcement, AI and IT tools the EU and its Member States could use to facilitate private enforcement of consumer protection legislation will not be excluded from the research.

---

<sup>7</sup> See p. 4-5 of the Terms of Reference (ToR).

**‘Consumer authorities’** is used in a broad sense and includes market surveillance authorities insofar as they are responsible for issues that are relevant for the protection of the rights and safety of consumers.

Central in this study is the term **‘tool’**. For the purpose of this study, we adopted a broad definition of this term. In particular, we define it to include:

- tools that are readily available for use by people with relatively little IT knowledge; this means that the tool has a user interface where the authorities staff members can insert e.g. websites to be analysed, searches to be carried out;
- tools that consist of a methodology that can be used to discover certain infringements or suspected practices and that often include computer code;
- tools that merely consist of computer code.

More specific concepts will be defined where appropriate.

## 1.3 Structure

The study consists of three main parts: inventory of AI/IT tools (Section 2), in-depth assessment of the tools (Section 3) and recommendations (Section 4).

In **Section 2**, we describe the process of creating an inventory of the tools. Specifically, we present what methods were implemented to identify the AI/IT tools that have been recently developed and/or are currently available on the market. We also discuss first results indicating the number of tools identified depending on the source of data, clustering them in helpful categories and giving a brief overview of the first findings (e.g. what types of tools are the most frequent).

**Section 3** proposes an exploratory framework for the assessment of such tools. Observations range from very general long-term directions that could be taken when choosing or developing AI/IT tools for monitoring and enforcement of consumer protection laws to more detailed suggestions as to specific issues that need to be taken into account when employing these tools in public interest (such as ethics, cybersecurity or accuracy).

**Section 4** discusses a series of recommendations based, tackling short-term enforcement suggestions, as well as long term implications of the features of the identified tools for the purpose of policy-making.



## 2 Inventory of AI / IT tools

### 2.1 Introduction

The first step in this study was to identify the tools that would fall under the definition of a tool and fulfil the criteria specified in section 1.2.

We aimed at including tools developed by either researchers, private entities or consumer authorities and that could potentially be relevant to this study. Since we wanted to make sure that our list is as comprehensive as possible, we used three different data collection methods.

We distributed questionnaires (see 2.2.1 for more details) to researchers as well as consumer authorities and organizations to find out what are the tools that they know. We also searched various online databases (2.2.2). Finally, we also included the tools that we learned about via (electronic) 'word of mouth' that is either by talking to other researchers and experts or through social media (see 2.2.3 for more details).

### 2.2 Methodology

#### 2.2.1 Questionnaires

Questionnaires were distributed to **21 consumer authorities and organizations** from the following countries: France, the UK, Finland, Estonia, Norway, Sweden, the Netherlands, Germany and Belgium (see Annex 4). These countries were selected either because of their high score on the Digital Economy and Society Index (2018)<sup>8</sup> or because of their recent activity or interests in developing or using AI monitoring tools in public services. We distributed further questionnaires to over a hundred researchers specializing in the field of consumer protection as well as computer and data science. The response rates were about 30% and 10% respectively.

We used two sets of questions (see Annex 2 and Annex 3). Consumer authorities and organizations were asked whether they conduct monitoring and investigations on infringements of consumer protection laws in online markets. We further inquired about the tools that are or could be used to carry out such activities. Finally, we also asked about the criteria that are considered relevant for quality assessment of these tools as well as about the most prominent issues faced by consumer agencies and organizations when carrying out online monitoring investigations. A shorter version of a questionnaire was sent to the selected researchers. Here, we explained the goal of the study and asked very broadly about any tools and resources that could be relevant. We additionally inquired about suggested quality assessment criteria.

The questionnaires were distributed and the responses collected using Qualtrics. Based on the questionnaires we learned about 64 unique tools.

---

<sup>8</sup> <https://ec.europa.eu/digital-single-market/en/news/digital-economy-and-society-index-2018-report>.

## 2.2.2 Database searches

Searches were carried out in legal databases (Jura, Jurisquare, Dalloz, HeinOnline, BeckOnline), a social sciences database (Web of Science), a general database for scholarly articles (Google scholar), and a number of software and patent registries.

Some of the databases are closed, meaning that they are only accessible against payment and do not allow for automatic extraction of information. This is the case for the legal databases and the social sciences database consulted. The other databases were open: freely accessible and automatically searchable. Therefore, we conducted two types of searches: manual and automatic.

In all databases, we ran 110 search strings (see Annex 5) that can be clustered into 11 categories. Some categories were later grouped together for analysis purposes finally resulting in 8 following categories: web evidence, advertising, dark patterns, geoblocking, scam (online fraud), unfair terms issues, price transparency issues, compliance.

In the closed databases, we manually verified the first 10 hits for their relevance and included the relevant hits into the project database. Based on this search strategy we identified 56 tools.

In the open databases, web scraping (i.e., automated extraction of information from the internet) was used based on the defined search strings. This automatic search resulted in a list of 3050 tools. Since this number of tools would not be possible to process, for further analysis we decided to draw a sample of 357 tools.

For an overview of the databases' search process, see Figure 1.

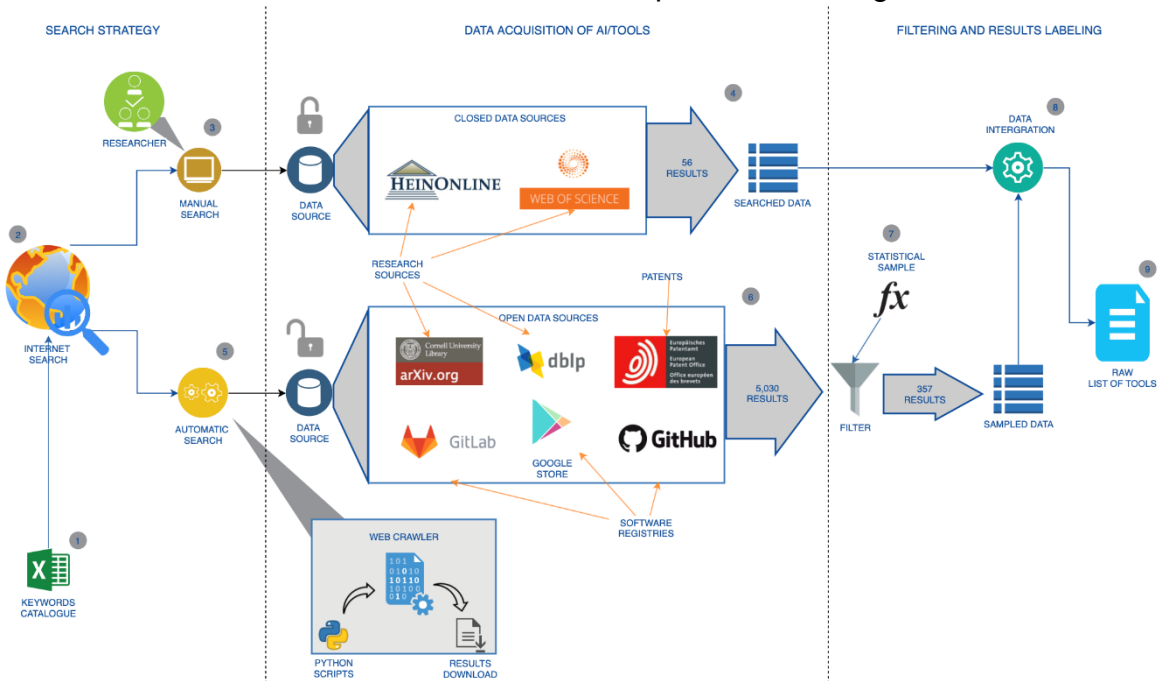


Figure 1: AI/IT Tools data acquisition workflow from different online sources

### 2.2.3 Word of mouth

The list compiled based on questionnaires and database searches was finalized by adding the tools that we had heard about when attending conferences but also systematically collecting information posted on social media by experts in the field. This way we learned about the 42 cutting-edge tools as well as those tools that could be adjusted or further developed to monitor the online market and enforce consumer protection laws.

## 2.3 Filtering out irrelevant tools and interim results

Based on the three steps described above, we compiled a list of 518 *potentially* relevant tools. To make sure that the list includes only tools that are indeed relevant to this study, we conducted an individual check. Each tool was coded independently by three members of our research group who indicated whether the tool is relevant or not. When assessing the relevance we considered whether the tool can serve the following purposes:

- detecting specific infringements of EU consumer law;
- detecting practices that are considered unfair by consumers and researchers; albeit not necessarily from an EU law or even legal perspective;
- helping authorities to better manage tasks such as complaints handling;
- helping authorities collect evidence of potential infringements.

In the case of tools mentioned in a response to our questionnaires, we also checked whether we could find more information about the tools concerned (e.g. on a website). A tool was included in the final list of relevant tools if it was coded as 'relevant' by all three researchers.

This resulted in a list of 73 tools. In the sections below, we briefly describe the tools on that list. In particular, we focus on the tools used by the authorities that responded to our questionnaire (see 2.2.1). We also briefly discuss different types of tools found in other sources

## 2.4 Interim conclusions

Most of the tools mentioned by the consumer authorities are of a rather general nature. Nevertheless, they are useful for the work carried out by the authorities, as shown by the fact that they are actually used. They help consumer authorities to collect evidence as well as to obtain and check information about suspicious websites. When selecting the tools we subjected to the detailed assessment, we, therefore, included tools that fulfil these purposes.

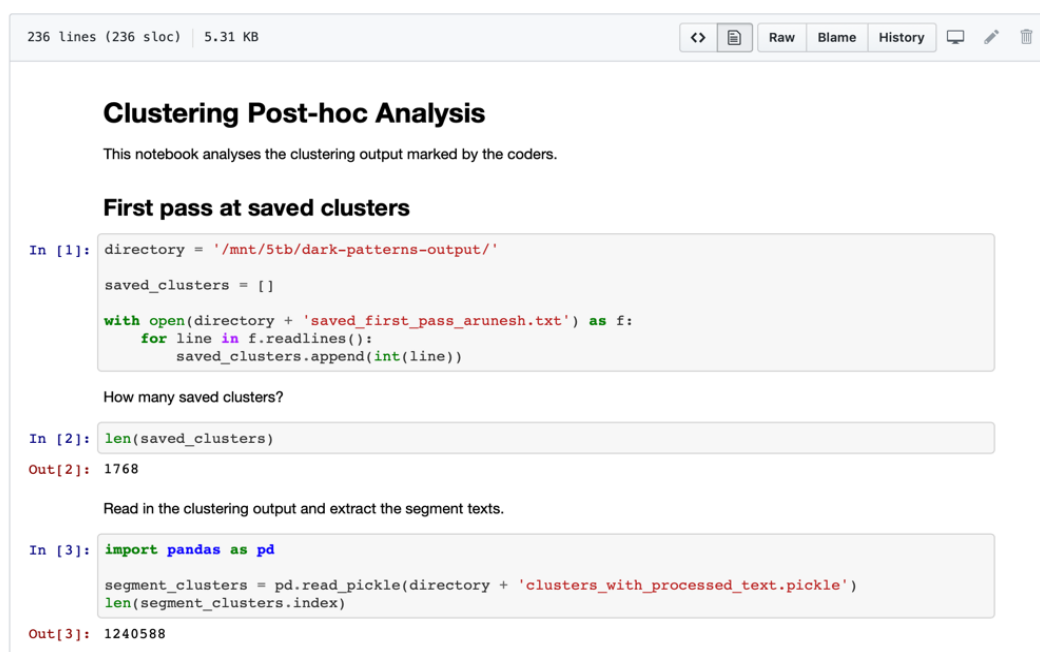
In contrast, tools found through database searches, word of mouth and that we learned about from researchers are more specifically targeted towards monitoring compliance with and detecting violations of consumer protection laws as well as providing consumers with tools to deal with issues they encounter when using online markets. The first ones are usually in a form of a computer code or a method, the latter come often with a user-friendly interface.

### 3 Proposed criteria for tool assessments

#### 3.1 Methodology: Selecting tools for further assessment

Before delving into the proposed framework for assessing tools which can be used for market monitoring, a few considerations need to be emphasized with respect to the general nature and goal of this assessment.

First of all, the tools collected for this report reflect **computer software**, or in other words, computer programs that run a set of instructions relating to the performance of particular tasks. This is a very broad definition that also includes executable programming code which may not have an interface. An interface is what makes a computer programme easy to use also by persons who do not have programming skills. A difference between executable programming code without and with an interface can be observed in Figure 2 and Figure 3 below.



```
236 lines (236 sloc) | 5.31 KB
<>  [ ]  Raw Blame History  [ ]  [ ]  [ ]

Clustering Post-hoc Analysis
This notebook analyses the clustering output marked by the coders.

First pass at saved clusters

In [1]: directory = '/mnt/5tb/dark-patterns-output/'
        saved_clusters = []
        with open(directory + 'saved_first_pass_arunesh.txt') as f:
            for line in f.readlines():
                saved_clusters.append(int(line))

How many saved clusters?

In [2]: len(saved_clusters)
Out[2]: 1768

Read in the clustering output and extract the segment texts.

In [3]: import pandas as pd
        segment_clusters = pd.read_pickle(directory + 'clusters_with_processed_text.pickle')
        len(segment_clusters.index)
Out[3]: 1240588
```

Figure 2: Dark Patterns study on Github – this programming code designed to identify dark patterns on shopping websites needs to be executed by a user with programming knowledge since it does not have an interface.

## CLAUDETTE

An Automated Detector of Potentially Unfair Clauses

Copy your text here

Submit

[About](#) [Cite](#) [Contact](#)

Figure 3: Claudette tool interface – the text of an unfair contract term can be copy-pasted and ran into the software by any user, regardless of their information literacy.

Secondly, the assessment of computer software is a **context-dependent task**. For instance, the ISO/IEC 25000 standards series (System and Software Quality Requirements and Evaluation) proposes a complex framework of evaluating software product quality. The product quality model is further defined in ISO/IEC 25010, and it comprises eight quality characteristics as shown in Figure 4 below. What these characteristics reveal, is that software assessment is a process heavily dependent on context. Features such as functional completeness, correctness or appropriateness mirror the specified tasks and objectives of the software product, which are in turn determined by the needs of the users. To put it differently, software needs to be built on user needs. Assessing tools which are built for a wide variety of user types, and dealing with an equally wide variety of legal issues poses serious limitations to the application of traditional software assessment principles. The most important effect of requiring software to be built on user needs is the practical impossibility of empirically testing functional correctness since for a lot of the tools, their accuracy cannot be detached from the data used to build it. Without having access to data, or data descriptions, such as may be the case with commercial products, it is impossible to fully assess accuracy.



Figure 4: ISO/IEC 25010 standard

Thirdly, software assessment is a **highly technical and specialized task**. Software testing is normally done by and for programmers. The translation of this testing into features used to describe software products for end-users or organizations commonly results in technical assessment reports. The goal of this study, however, is not to make an overview for a technical audience, but rather to describe technical information for consumer enforcement professionals.

Lastly, **software development** has seen a lot of **innovation** in the past decade. Both from a legal and from a business perspective, the software can be distributed using two models: **Software as a Product (SaaP)** and **Software as a Service (SaaS)**. Unlike SaaP, which entails the purchasing of a licence to install software on local computers, SaaS is 'a delivery and licensing model in which software is accessed on the web via a subscription',<sup>9</sup> through the use of cloud computing. SaaS is increasingly adopted as a business solution<sup>10</sup> and it has clear advantages such as lower costs,

<sup>9</sup> Cisco, 'What Is Software as a Service', available at

<https://www.cisco.com/c/en/us/products/software/what-is-software-as-a-service-saas.html>.

<sup>10</sup> IDG, 'Cloud computing survey' (IDG Communications, 2018), available at

<https://cdn2.hubspot.net/hubfs/1624046/2018%20Cloud%20Computing%20Executive%20Summary.pdf>.

scalability, upgrade management or ease of use,<sup>11</sup> it may pose considerable risks such as performance/data control or security.<sup>12</sup> This affects aspects of software management which may be important for an organization in the maintenance, security or portability of software products.

With these considerations in mind, this section provides an assessment based on original criteria developed on the basis of the ISO/IEC 25000, but adapted to the goal of this project, namely to find and describe AI/IT tools which can be used for market monitoring and/or for the enforcement of consumer law. To make comparisons easier, a lot of the criteria identified below use labels, which are further elaborated upon where necessary. In addition to these characteristics, **maintainability** and **security** are factors that have been taken into consideration when assessing the tools, and are overall evaluated in Section 3. Moreover, **privacy** and **ethics** are considered independently under the same section, as broader policy and information management considerations.

An exploratory, general framework which could further assess IT / AI Tools which may be deployed by consumer protection authorities could focus on the following characteristics.

**a. Name of the tool**

Used to identify the tool.

**b. Website**

The website where the tool can be found online.

**c. Description**

This part generally describes what the tool does, and how it works, based on product specifications and of a test of the tool (where possible). Where possible, the basic functions of the tools have been tested at least once.

**d. Consumer protection issue addressed**

This part is based on the classification made in section 2.2.2. Seven categories<sup>13</sup> of legal issues can be acknowledged:

- *Unfair contract terms*
- *Unfair commercial practices (e.g. advertising)*
- *Price transparency*
- *Geoblocking*
- *Scam (online fraud, e.g. phishing)*
- *Consumer forensics*

---

<sup>11</sup> M. Sylos, 'Top five advantages of software as a service (SaaS)' (IBM, 18 September 2013), available at <https://www.ibm.com/blogs/cloud-computing/2013/09/18/top-five-advantages-of-software-as-a-service-saas/>.

<sup>12</sup> M. Jansen, A. Joha, 'Adoption Challenges of Introducing and Implementing SaaS', in Mahmud Akhter Shareef, Norm Archer, Yogesh K. Dwivedi, Transformational Government Through EGov Practice: Socio-Economic, Cultural, and Technological Issues (Emerald Publishing, 2012), p. 241.

<sup>13</sup> The classification made in section 2.2.2 included also one additional category, i.e., web evidence. This category was excluded from this list, since here we focus only on legal issues.

These issues may sometimes overlap (e.g. unfair contract terms, taken together and applied systematically by companies, can also constitute unfair commercial practices). However, since some of the tools put emphasis on specific elements of the consumer protection instruments in force at the European level, we decided to treat them as separate categories. The first four categories deal with legal issues specific to the consumer *acquis*. The latter two categories ('scam' and 'consumer forensics') entail, in addition to elements of consumer protection, aspects of criminal and procedural law. In particular, in this report, consumer forensics is a concept used to refer to a wide range of scientific approaches to the procedural need to investigate and gather evidence regarding fraudulent practices impacting consumer rights.

#### **e. Developer**

Identifying the developer entails, on the one hand, specifying the name of the person or entity who developed the tool, and on the other hand which developer category (s)he/it falls under. We used the following categories:

- *Commercial*: companies or industry entities that develop / use software for commercial or industry-wide purposes.
- *Public administration*: European or national authorities tasked with the enforcement of European law.
- *Academia*: research groups developing academic data gathering and / or analysis methods as well as prototypes.
- *Individual developers*: physical persons creating open-source software that is available on the Internet, commonly for free.

This characteristic is particularly important because it is linked to maintainability. Commercial solutions will almost always entail that the company will maintain a specific tool (unless it decides to stop supporting it, which is not a problem for SaaS models, where the decision to stop supporting a tool will entail pulling the tool off the market by not offering it as a service any longer). It must also be emphasized that a lot of tools are made online by programmers either in a voluntary manner, or as part of academic research. We distinguish between the context in which these tools are made: the 'academia' category entails code that is published as a result of an academic project, whereas the 'individual developers' category is broader and includes any tools developed on the Internet and not affiliated to an academic project or a company. In the case of both these categories, maintainability will not be done by the respective developers, but needs to be done by the persons who decide to use these tools.

#### **f. Programming language used**

Any tool or software program is conceived by a collection of technologies and programming languages during its development. Here, we will be focusing on registering two types of programming languages: a) Programming languages used to develop the tool user interface (if applies) and b) Programming languages used to develop the core functionality of the tool. The latter is mandatory for any software for



full functionality, on the other hand, a user interface is not mandatory for a tool to function.

#### **g. Price**

Where the price of a tool (or lack thereof) is known, it will be mentioned in the assessment. From the perspective of price, tools fall within various categories:

- *Paid (known price)*: the tool is available as is, for an identifiable price. This is mainly the Software as a Product (SaaP) model, where a user can purchase a license to use the tool. Some software subscription services (Software as a Service – SaaS model) also have identifiable prices. An example in this respect is VPN services.
- *Paid (bespoke price)*: there is an available tool, however, for the tool to be scaled, or to be fitted to specific enforcement objectives, further development is needed. It must be noted that it is practically impossible to ask for quotes for such development services without clearly specifying the goals and tasks which the software ought to perform. Once more, this must reflect the needs of end-users (e.g. national consumer authorities), which vary greatly.
- *Free (open source)*
- *A free version* (with limitations in terms of functionalities, number of searches, users etc.)

#### **h. Whether the tool is already used by consumer authorities**

This characteristic is meant to reveal whether a given tool is already in use by a national consumer authority.

#### **i. Languages**

Tools are mostly made for specific languages, which are identified in the assessment. Scalability in terms of additional languages can also be discussed under this heading.

#### **j. Type**

Monitoring can be defined in different ways. For the purpose of this study, monitoring is considered to be any scrutiny process which can be done at scale, namely on the basis of a large number of observations (e.g. in the thousands, tens of thousands, millions, etc.). Individual instances of scrutiny are referred to as investigations, and where it is necessary/possible to record these observations, the study speaks about 'evidence tools'.

Four different types of consumer tools are identified:

- *Market monitoring tools*: tools that can monitor the compliance with consumer protection rules (e.g. unfair terms or unfair commercial practices). These tools operate on a large scale (e.g. at the national level) and entail big data analysis (e.g. crawling tens/hundreds of thousands of webshops). These tools may be used by various actors, such as consumer organizations, data protection organizations, but also businesses.
- *Investigation tools*: tools that allow for individual investigations into specific issues (e.g. finding the IP address of a given trader). Those tools operate on

a small scale and are suitable for in-depth investigations into specific actors or practices.

- *Evidence tools*: tools that record the findings of large or small scale monitoring/investigation (e.g. saving screenshots of webshops).
- *Consumer empowerment tools*: tools that are made for consumers as end-users. Although not primarily the focus of this study, some consumer empowerment tools have great scalability potential, specifically identified in the characteristic 'scalability' as explained below under l).

#### **k. Technology**

This part is used to briefly describe what technology the tool uses, wherever this information is available. A distinction is made between tools using AI, and tools not using AI. This distinction is important due to the fact that some types of machine learning may pose additional privacy issues which consumer protection agencies need to be aware of. A general overview regarding privacy questions is available under section 4.3.3.

#### **l. Accuracy**

As mentioned above, measuring software accuracy is a highly complex process. For this study, this criterion is used to indicate whether any perceivable issues can be deduced from software configuration. When the developer reports the accuracy of the tool, the 'self-reported' category is indicated.

#### **m. Scalability**

Linked to the characteristic 'type', scalability is presented as an indication of which tools can be made into a different type, namely which e.g. consumer empowerment tools can become monitoring tools, etc.

#### **n. Ease of use**

Different operationalizations are possible. Ideally, 'Ease of use' means it can be directly implemented by enforcement agencies. However, since implementation requires processes, procedures, permissions, and because ease of use depends on the precise purpose and internal organizations, none of the tools can be implemented directly. In this study, ease of use is analysed by means of tools with and without user interface. In addition, other observations regarding ease of use will be made. Since the tools under scrutiny may or may not have interfaces, the programming skills necessary to deploy them varies significantly. This may considerably affect the way in which a national consumer protection agency may use them, whether available for a specific purpose already, or scaled/modified for other purposes. On the basis of ease of use, two categories are distinguished:

- *Tools with an interface*: these tools are seemingly easy to use by anyone, even users without any programming knowledge, as they can be installed and run easily. The question of whether the interface is sophisticated enough, and in itself, user-friendly, is a separate question, addressed in the table where relevant.

- *Tools without interface*: these tools mostly consist of executable programming code that is freely available on the internet as part of various public and/or private research projects, and which requires programming skills to operate.

## 4 Recommendations

The question as to which tools to use or to choose cannot be answered in isolation. It has to be assessed in a broader context, which is further elaborated below.

### 4.1 Reframing the need for technological solutions

Technological innovation is constantly challenging the efficiency and effectiveness of law enforcement. In recent times, the pace of this innovation has been increasing at record rates. Since this phenomenon already led to new markets, actors, organizational structures, business models, but also to new types and scales of consumer harm, public administration agencies from around the world are currently struggling to keep up with this pace. The tools identified in this study depict the current state of affairs in terms of tool development. Focusing on how the tools ought to be used, shared and implemented by consumer authorities can constitute a very important short-term focus. However, it is also necessary to reflect on how a broader strategy could look like in the long term. Public administration is called upon to use the same forces of innovation to gain new monitoring and enforcement competencies, and thus be able to deal with the increasing level of sophistication of the technologies used by (malicious) market actors. This has been, for instance, articulated in the **public interest technology** approach proposed by public policy scholars and professionals from the Harvard Kennedy School of Government.<sup>14</sup> Defined as ‘the study and application of technology expertise to advance the public interest / generate public benefits / promote the public good’, public interest technology may take many different shapes, depending on the models employed in the development of tools:

- **In-house expertise:** public authorities developing in-house technology expertise (e.g. the data units of the UK CMA & ICO).

Benefits	Risks
Can tackle very specific needs and priorities	Unsustainable financial investments for infrastructure and human resources
Control over security and privacy	Serious expertise for maintenance, privacy & security is required
Control over ethical framework	Development costs may be higher (depending on the tool)
Heavily increasing organizational information literacy	Developing legacy systems that lose flexibility

---

<sup>14</sup> B. Schneier, ‘Public-Interest Technology Resources’ (Schneier Blog, 18 November 2019), available at <https://public-interest-tech.com>.

- **Outsourcing:** public authorities outsourcing bespoke technological solutions to market actors (mostly offered as SaaS).

The vast majority of market solutions we investigated are SaaS (e.g. DomainTools, Distill.io, Versionista). The question of whether SaaS tools ought to be built in-house or outsourced to market actors is by no means new. However, purchasing licenses for Office 365 is not the same as having a bespoke fraud detection tool. SaaS entails that the provider makes available absolutely all the necessary infrastructure - public authorities need not worry about not having the necessary staff members or the necessary equipment to develop these solutions.

Benefits	Risks
Running services externally, on the cloud → no need for in-house installation and maintenance of software	Commercial reliability and long-term transacting
No need for a data unit / human resource management and the required financial resources	Risk of lock-in
No need for financial investment in computing power	Limited customisation
Market competition entails better products and state of the art innovation	No/limited in-house (for consumer organisations) technical expertise to align contractual expectations
SaaS providers can specialise on a specific area of expertise	no/limited control over data ownership, privacy and cybersecurity
Backup and recovery ensured by SaaS provider	global supply chain leading to different legal standards applicable
	loss of data in case of SaaS provider bankruptcy
	no/limited accuracy control

- **Grassroots:** civil society developing consumer empowerment tools independently from the mandate of public authorities.

Benefits	Risks
Can complement solutions by public authorities	Sporadic development due to non-systematic funding

Citizen empowerment	No/limited accuracy control
Information literacy for citizens	Potential misalignment with public authority priorities
Mainly non-political/non-commercial	No/limited control over data ownership, privacy and cybersecurity
	Potential long-term maintenance issues

- **Public industry tools:** industry organisations developing consumer empowerment/market tools independently from the mandate of public authorities, to help correct/prevent market failures;

Benefits	Risks
Can complement solutions by public authorities	No/limited accuracy control
Citizen empowerment	Potential misalignment with public authority priorities
Information literacy for citizens	No/limited control over data ownership, privacy and cybersecurity
Mainly non-commercial	Underlying commercial/political interests

- **Academic research:** independent research institutions (e.g. universities and research centres) developing academic tools to be used by public authorities and/or consumers.

Benefits	Risks
Valorising the power of academia as public repositories of knowledge	Project-based collaborations, no follow-up
State of the art research expertise	Determining necessary expertise due to no / limited in-house technical experience
Interdisciplinary centres and networks (law & computer science)	No centralization of tools
Potential for strategic long-term coordination/centralisation	Difficulty with necessary interdisciplinary collaborations
Bespoke tools based on the needs of public authorities	Potential misalignment of public policy in practice and academic neutrality

The five approaches to public interest technology enumerated above have various risks and benefits, which can be identified on different levels such as strategic, organizational, political, technical or economic. When considering these risks and

benefits, however, there are three aspects of this choice which need to be emphasized: **data ownership and privacy**, **ethics**, and **cybersecurity**.

Issues relating to **data ownership and privacy** arise primarily out of the combination of malicious and / or commercial interests and the business models of the new data economy, which remain largely unexplored. For instance, companies active in the industry of data enrichment operate without any visibility, and only become known if the media covers situations of data breaches in which they are involved<sup>15</sup>. Operating in grey areas where their activities are not immediately unlawful (e.g. the mere collection of email addresses from public websites may be permitted in different jurisdictions)<sup>16</sup>, data brokers have created incentives for companies that operate more visibly on digital markets to monetize user information. In an extensive case study on over 100 VPN service providers, review and comparison service 'The Best VPN' fact-checked privacy policies upon login, and found that 26 out of the 117 tested services retrieve and keep user information such as bandwidth, timestamps, IP address and browsing history (see Figure 5).

---

<sup>15</sup> Data enrichment or augmentation is defined as '[t]he process of enhancing existing information by supplementing missing or incomplete data. Typically, data enrichment is achieved by using external data sources.' In other words, if companies want to gather more information on their clients, whether legal persons or natural persons, they can hire other companies that try to collect more data points from external sources (e.g. media, social media, public registries, etc.), M. Allen and D. Cervo, 'Data Quality Management' in Multi-Domain Master Data Management Advanced MDM and Data Governance in Practice 2015, p. 131-160.

<sup>16</sup> See the People Data Lab data breach <https://www.troyhunt.com/data-enrichment-people-data-labs-and-another-622m-email-addresses/>

VPN Provider	Type of Logs	Privacy Policy	Share/Sell Data?
1. McAfee Safe Connect 1+ million installs	<ul style="list-style-type: none"> <li>Websites you visit</li> <li>Your IP address</li> <li>Apps installed on your device</li> <li>Timestamps</li> </ul>	We may collect and use... details about your internet or network usage (including URLs or domain names of websites you visit, information about the applications installed on your device, or traffic data... <a href="#">(source)</a>	We may also share aggregate data that does not identify you or any specific device with third parties.
2. Hola VPN 50+ million installs	<ul style="list-style-type: none"> <li>Websites you visit</li> <li>Time on those pages</li> <li>Timestamps</li> </ul>	We collect log data that may include browser type, web pages you visit, time on those pages, access items and dates... <a href="#">(source)</a>	We may also transfer or disclose Personal Information to our subsidiaries, affiliated companies.
3. Hoxx VPN <1 million installs	<ul style="list-style-type: none"> <li>Websites you visit</li> <li>Your IP address</li> <li>Timestamps</li> </ul>	We collect log information about ... access times, pages viewed, your IP address... <a href="#">(source)</a>	We may share your data with our services providers... We have contracts with our service providers that prohibit them from sharing the information about you that they collect or that we provide to them with anyone else, or using it for other purposes.
4. Hotspot VPN 1+ million installs	<ul style="list-style-type: none"> <li>Websites you visit</li> <li>Things you type into fields</li> <li>Your device information</li> <li>Your IP address</li> </ul>	We log information and other data from your device, such as webpage addresses and data fields... <a href="#">(source)</a>	We may share, sell, transmit, or otherwise make available to third parties information that does not include personally identifying information.

Figure 5: The Best VPN study<sup>17</sup>

If consumer protection authorities use free VPN services, or paid services which collect user data and engage in third-party transactions with respect to this data, it may very well be that an entire market of profiling public authorities activities can develop. For this reason, it is of utmost importance that when choosing what tools to work with/develop, consumer authorities carefully consider how much control they have over the data they produce, and whether there are any risks for this data to be passed on to third parties.

Regarding data protection, building and/or using tools, consumer authorities need to make sure to respect the GDPR: Insofar as personal data within the meaning of the GDPR are used to train AI tools, and/or insofar the use of AI/IT tools results in the processing of personal data wholly or partly by automated means or in the processing other than by automated means of personal data which form part of a filing

<sup>17</sup> R. Mardisalu, '100+ VPN Logging Policies Debunked', <https://thebestvpn.com/118-vpns-logging-policy/>. Express VPN only keeps logs on one of the investigated data categories (bandwidth).



system or are intended to form part of a filing system, the GDPR will of course need to be respected.

With respect to **ethics**, investigating certain companies against which complaints were made, or which are suspected to engage in harmful activities against consumers is legitimized by the legal mandate of consumer authorities at the national and European level. Monitoring entire industries requires the collection and processing of considerable amounts of data from a variety of sources. This potential power, without a transparent mandate and accountability, may be construed as state surveillance.<sup>18</sup> However, tracking commercial partners<sup>19</sup> should not pose the same ethical questions as profiling individuals. This is all the more so because Article 29 of the CPC Regulation provides a legal basis for sweeps, ‘concerted investigations of consumer markets through simultaneous coordinated control actions to check compliance with, or to detect infringements of, Union laws that protect consumers’<sup>20</sup>. Recital 9 CPC Regulation which elaborates on sweeps, states that ‘authorities should have access to any relevant documents, data and information that relate to the subject matter of an investigation or concerted investigations of a consumer market (‘sweeps’) in order to determine whether an infringement of Union laws that protect consumers’ interests has occurred or is occurring, and in particular to identify the trader responsible, irrespective of who possesses the documents, data or information in question, and regardless of their form or format, their storage medium, or the place where they are stored. Competent authorities should be able to directly request that third parties in the digital value chain provide any relevant evidence, data and information in accordance with Directive 2000/31/EC of the European Parliament and of the Council (4) and in accordance with the legislation on personal data protection’. This recital seems to allow consumer authorities to use web scraping when carrying out sweeps, even when the terms and conditions of the platform on which the trader is active prevent this.<sup>21</sup>

---

<sup>18</sup> Dutch courts have held that where personal identity is replaced by an ID number, there is no impairment of privacy when this data is scraped, as scraping entails gathering ‘publicly accessible information, that is data visible to everyone; this also applies to the ‘listing ID’.

<sup>19</sup> In any case where they are legal persons. If a physical person is acting in a professional and commercial capacity in his own name things may be more complicated. For example, the GDPR applies. We also draw the attention to a recent judgment from a Dutch court that addressed the limitations of state authority in profiling citizens to prevent social security fraud (Syri; ECLI:NL:RBDHA:2020:865).

<sup>20</sup> Art. 3 (16) CPC Regulation.

<sup>21</sup> Within the framework of sweeps and absent a clear legal framework for web scraping, national courts were faced with questions about the legality of web scraping from the perspective of either data protection or tort law. Even when websites indicate in their terms and conditions that no scraping is allowed, the Den Haag Appeals Court, in applying Irish law, held that if there is no database right or copyright protecting information that is publicly available, automated means to collect data are not unlawful (ECLI:NL:GHDHA:2018:61, para 79). In US law, trespassing, as a tort law concept, has been used to analyse situations where commercial scraping may damage the business interests of a platform, but even in these cases, to the extent that the company undertaking scraping, and the company whose information is scraped, are not direct competitors, no tort law remedy is available (see for instance QVC Inc. v. Resultly LLC, 2015 WL 1187500 (E.D. Pa. March 13, 2015)).

Moreover, tools using AI need to respect the (forthcoming) ethical framework: There is wide consensus that AI needs to respect certain ethical principles. In June 2018, the European Commission set up a High-Level Expert Group on Artificial Intelligence (AI HLEG) that was given the task to draft Ethical Guidelines for Trustworthy AI. The Expert Group published the Guidelines drafted in April 2019.<sup>22</sup> At the start of her term as President of the European Commission, Von der Leyen promised AI legislation within 100 days.<sup>23</sup> In February 2020, a white paper on ethical guidelines for AI is expected. If the Commissioner succeeds in her aim, this will rapidly be followed by a legislative proposal.

With respect to **cybersecurity**, in performing digital monitoring and investigations through the use of tools built or purchased for the purposes of enforcing consumer protection, it is of utmost importance to consider informational security from two main perspectives.

First, public authorities need to be aware that cyberattacks are becoming more prevalent, and need to take measures to protect themselves. The first US nation-wide survey regarding local government and cybersecurity paints a grim picture, where, on the one hand, 27.7% of the institutions surveyed reported they were able to detect cyberattacks occurring hourly, and on the other hand almost 30% of the respondents were not able to report whether they were being attacked.<sup>24</sup> All tools will be faced with some infrastructural risks, depending on their nature. The highest performing tools from a cybersecurity perspective are those that have access to the most data and / or have the highest industry expertise (e.g. DomainTools). However, it is difficult to assess what kind of risks arise within these transactions in terms of information security (e.g. if such companies deal with national security threats, it is likely to assume that various national intelligence services will be monitoring/controlling, as the case may be, some of these activities). In the case of open source tools (e.g. tools made by unknown developers), reverse engineering may pose problems, which is why further tailoring of these tools ought to be considered.

Second, consumer law enforcement often overlaps with criminal law. For instance, scammers who engage in robocalls to deceive consumers into concluding contracts related to IT services use spyware or other types of malware to affect the integrity of data on consumers' personal computers, in order to persuade them to acquire IT services. Forensic investigations into the type of malware used or its provenance (e.g. sold on dark web marketplaces) require considerable cybersecurity expertise, as well as coordination with other law enforcement authorities (e.g. police). This will be further explored in section 4.3.

---

<sup>22</sup> High-Level Expert Group on AI, Ethics Guidelines for Trustworthy Artificial Intelligence, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

<sup>23</sup> U. von der Leyen, Political guidelines for the next European Commission 2019-2024, [https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission\\_en.pdf](https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf).

<sup>24</sup> D.F. Norris et al., 'Cyberattacks at the Grass Roots: American Local Governments and the Need for High Levels of Cybersecurity' (2019) 79(6) Public Administration Review 895.

## 4.2 Invest in cooperation and information sharing with other consumer authorities and in the tools that allow this to take place in an efficient way

In a digital environment, national borders do not exist. Websites and webshops of a certain Member State or even a non-EU-country are generally accessible throughout the EU so that consumers throughout the EU are potential victims of unfair practices used on those websites and/or by online businesses during the stage of the performance of the contract. Given the fact that the main part of the Member States' consumer law is based on EU law that is increasingly based on maximum harmonisation, a set of facts that amounts to an infringement of the consumer law of a certain Member State is highly likely to also be infringement of the consumer law of other EU Member States.

If one consumer authority is aware of such an unfair practice it is therefore important to communicate this information to other consumer authorities in order for them to take appropriate measures. In order to enable efficient communication between consumer authorities they should use the same system to qualify and classify complaints or other notifications of infringements. It could be considered to use the same case management system, including the different steps that a certain authority has taken and the results thereof. Since most infringements are (indirectly) based on EU directives or regulations anyway, it makes sense to determine the labels according to the relevant EU law instruments and provisions thereof. It could be investigated to what extent consumer complaints contain sufficient information to enable automatic labelling. In case the EU law instrument that is at the basis of the rule of consumer protection is a Directive, implemented in national legislation, the national provision infringement can be mentioned in the file. In cases where the precise legal nature of the potential infringement is not yet defined, one or broader residuary categories could be used.

The **Themis project** developed a tool that aims specifically at the exchange of information between Member States and the Commission during the phase of implementation of European legislation and in case of infringement proceedings by the Commission against a Member State. It facilitates the notification of national measures transposing EU directives, the communication between the Commission and the Member States before launching an infringement procedure (EU Pilot) and the communication of replies and requests for extending the deadlines to infringement decisions.<sup>25</sup>

A system allowing the communication between the Member States' consumer authorities and the Commission regarding national enforcement actions of EU Regulations or national legislation transposing EU Directives seems to be a logical **extension** of the Themis project.

---

<sup>25</sup> [https://ec.europa.eu/isa2/actions/information-exchange-between-eu-and-member-states-monitor-application-eu-law\\_en](https://ec.europa.eu/isa2/actions/information-exchange-between-eu-and-member-states-monitor-application-eu-law_en).

### 4.3 Invest in cooperation and information sharing with criminal law, competition law and data protection enforcement agencies and discuss access to some of their AI/IT tools

As explained below, infringements of consumer law in the digital sphere are often closely related to cybercrime, infringements of competition law and infringements of data protection law. The sharing of information and the bundling of competences in these fields is therefore of utmost importance. This aligns with Van der Leyen's plan to create a joint Cyber Unit to speed up information sharing and increase the level of protection against illegal behaviour in the online world.<sup>26</sup>

#### 4.3.1 Consumer law and criminal law

Some consumer law infringements or broader issues of consumer protection are at the same time criminal acts or at least closely related to criminal acts and even organized crime. This is, for example, the case for identity theft, sale of fake / inexistent goods (where the party pretending to be a seller never had the intention to actually deliver a product after having received payment), obtaining personal data in illegitimate ways to allow for personalised advertising, often aimed at the most vulnerable categories of consumers (e.g. elderly seeking good investments without any financial knowledge, people with terminal illnesses or people that already accumulated so much debt that they can no longer obtain credit from legitimate credit institutions).

For this reason close cooperation and (within the boundaries of privacy protection<sup>27</sup>) information sharing with criminal law enforcement agencies is important. A number of Horizon 2020 programmes<sup>28</sup> and ISF-Police programmes<sup>29</sup> invested in the development of tools facilitating the work of criminal law enforcement agencies. A number of these tools would also be useful or could be amended for use by consumer authorities.

---

<sup>26</sup> U. von der Leyen, Political guidelines for the next European Commission 2019-2024, [https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission\\_en.pdf](https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf).

<sup>27</sup> Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA, OJ L 119, 4.5.2016, p. 89.

<sup>28</sup> See for example <https://ec.europa.eu/info/funding-tenders/opportunities/portal/screen/opportunities/topic-details/su-fct02-2018-2019-2020;freeTextSearchKeyword=;typeCodes=1;statusCodes=31094501,31094502;programCode=H2020;programDivisionCode=31048010;focusAreaCode=null;crossCuttingPriorityCode=null;callCode=Default;sortQuery=openingDate;orderBy=asc;onlyTenders=false;topicListKey=topicSearchTablePageState>.

<sup>29</sup> [https://ec.europa.eu/home-affairs/financing/fundings/security-and-safeguarding-liberties/internal-security-fund-police/union-actions\\_en](https://ec.europa.eu/home-affairs/financing/fundings/security-and-safeguarding-liberties/internal-security-fund-police/union-actions_en).

This is, for example, the case for **FoBro** (Forensic Browser) developed by the UCD Centre for Cybersecurity & Cybercrime Investigation.<sup>30</sup> We have not been able to test the tool, but according to the developers, the aim of the tool is 'to gather and preserve online evidence, as seen during online investigations. Besides the capabilities to download videos from hundreds of portals, an automated plugin to secure Facebook profiles and a one-click Tor solution to enter the DarkNet, the tool provides full, unencrypted network traffic capture, which can be replayed post-investigation and provides additional support for viewing and capturing evidential material. This includes text searching, mouse-over information, automated scrolling, dynamic offline HTML with inline/offline media support, etc.'. This tool is not commercially available, and access is restricted to certain groups. According to the developers, consumer authorities would be eligible to obtain access to the tool.

Another tool that is developed by the same research centre and within the same framework that could be useful for consumer authorities is **DeepThought**. Again, we were not able to test the tool, and only obtained the following description: 'DeepThought conducts an examination of many of the artefacts contained on digital media and focuses on the retrieval of images, movies, documents, web history, emails, chat, keyword searching and registry from allocated deleted and unallocated space. It is configured to have a much faster completion time for the analysis of digital media, compared to that of a full forensic examination, whilst overcoming the inherent weaknesses that exist with standard triage and other commercial tools. Tests have found that DeepThought out-performs its commercial counterparts, such as EnCase, at image retrieval across multiple file-systems'.

In order to facilitate communication and information sharing between consumer authorities and criminal investigation teams and courts, it would be useful to develop a list of case patterns, functional descriptions of facts in non-legal terms that can give rise to both an infringement of consumer law and an infringement of criminal law in the legal systems of the member states. Except for a number of typical cybercrimes for which there is harmonised legislation,<sup>31</sup> criminal law is still mainly national law. Given the fact that the criminal legal systems of the member states have their own

---

<sup>30</sup> [http://www.ucd.ie/cci/projects/current\\_projects/freetool2.html](http://www.ucd.ie/cci/projects/current_projects/freetool2.html)

<sup>31</sup> European Parliament and the Council of the European Union, Directive 2013/40/EU of 12 August 2013 on attacks against information systems and replacing Council Framework Decision 2005/222/JHA, OJ L 218, 14 August 2013, pp. 8-14. <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0040&from=EN>. The articles of this Directive concerning crimes that may be related to consumer law infringements are: *Article 5*: Illegal data interference (deleting, damaging, deteriorating, altering or suppressing computer data on an information system, or rendering such data inaccessible, intentionally and without right, is punishable as a criminal offence, at least for cases which are not minor) and *Article 6*: Illegal interception (intercepting, by technical means, non-public transmissions of computer data to, from or within an information system, including electromagnetic emissions from an information system carrying such computer data, intentionally and without right, is punishable as a criminal offence, at least for cases which are not minor).

terminology based on their own legal tradition, the factual, functional description should also be linked to the corresponding national criminal law concepts.

#### 4.3.2 Consumer law and competition law

Many issues emerging in online markets, such as algorithmic pricing, unfair pricing or other unfair contract conditions that can be imposed as the result of the network effects, raise issues of both consumer protection and the protection of free competition. Cooperation between consumer and competition authorities in terms of information sharing and joint capacity building is likely to benefit the enforcement activities of both types of enforcers. In order to facilitate cooperation between the authorities, agreements may be needed in order to make sure that data collected by one authority can easily be processed by the technology used by the other authority.

#### 4.3.3 Consumer law and data protection law

Data obtained by means of infringements of the GDPR may be used as part of an unfair commercial practice, such as personalised advertising or personalised pricing aimed at taking advantage of consumers' personal weaknesses. Equally, when purchasing online goods and services consumers may inadvertently volunteer personal data and consent to the use of thereof for purposes way beyond what is necessary for the performance of the contract concluded. Close cooperation between consumer authorities and data protection authorities will help to see the links between infringements of consumer law and of data protection law and will prevent that the same investigation is carried out by both authorities operating in isolation. Here too the interoperability issue is to be taken into account.

#### 4.3.4 Cooperate with national and EU consumer associations

Consumer associations receive complaints from consumers. It would be helpful if they classified the complaints received according to the same classifier system as the EU and if they reported to the national consumer authorities and the EU about the practices and suppliers most complained of so that this information can be taken into account when determining EU wide or national enforcement priorities which can also determine the kind of detection tools that need to be developed with the highest priority.

Consumer associations may also help to improve the quality of tools such as **Scamdok** or **Sentinel** that are at least partially based on crowd sourced information about scams by inputting this information into the tools concerned and by including links on their websites to the tools concerned or by otherwise bringing these tools to the attention of their members so that the consumers themselves feed the tools with information about the tools and use the tools to prevent dealing with scammers.

## 4.4 Look into the possible use of (general) tools for policy development for the development of consumer policy

The tools we analysed in detail mainly focus on the **enforcement** of consumer law. However, some of the tools also provide insight in practices used by online suppliers that are considered unfair by consumers although they are not yet explicitly prohibited or even mentioned in EU consumer law instruments. This is for example the case for tools such as **Scamdoc** or **Sentinel** that partly rely on information supplied by consumers about what they consider as a scam or as being unfairly treated. Also, the **Dark patterns study**<sup>32</sup> allows to discover clusters of practices used on e-commerce sites. These clusters need to be subsequently analysed by a lawyer. Some of the clusters will concern practices that are already prohibited by consumer law, other clusters will concern practices that are completely harmless. The most interesting ones are those legislators or authorities were not yet aware of, but which on closer inspection seem to be misleading, manipulating or otherwise harming consumers. Such findings may provide inspiration for the **further development of consumer law**. A tool that aims specifically at policy development is **Sense4Us**.<sup>33</sup> We did not include this tool in the final list of tools, because it was of an entirely different nature than the other tools listed. Nevertheless, if there is a specific interest in the use of tools for policy development, this tool deserves further attention. More specifically, Sense4Us provides a prototype for 'policy modelling and simulation based on the information gathered from various online sources, including linked open data search results (...) and evidence extracted from online public political discussions'.<sup>34</sup>

## 4.5 Use the shortlisted monitoring tools in order to obtain a first indication of possible infringements

The most impressive monitoring tools in the list are the academic tools developed by Princeton University to investigate **dark patterns**, **affiliate marketing** and **web transparency (OpenWPM)**. These tools however entail further tailoring by software developers, so they can be used immediately in two circumstances: (a) if consumer authorities have data units or in-house programming expertise; or (b) if consumer

---

<sup>32</sup> A. Mathur et al. (2019) Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. Proc. ACM Hum.-Comput. Interact. 3, CSCW, Article 81 (November 2019), 32 pages. DOI: <https://doi.org/10.1145/3359183>. See also <https://webtransparency.cs.princeton.edu/dark-patterns/>.

<sup>33</sup> Project website: <http://www.sense4us.eu/>. Link to the tool: <http://dev1.egovlab.eu:4001/>. Publication explaining the tool: O. Ibrahamin and A. Larsson (2016) 'A systems tool for structuring public policy problems and design of policy options', Int. J. Electronic Governance, Vol. 9, Nos. 1/2, 4-26.

<sup>34</sup> O. Ibrahamin and A. Larsson (2015) Policy Modelling and Simulation Tool, [http://www.sense4us.eu/images/reports/611242-Sense4us\\_D6-2-Policy\\_Modelling\\_Tool-FINAL.pdf](http://www.sense4us.eu/images/reports/611242-Sense4us_D6-2-Policy_Modelling_Tool-FINAL.pdf), p. 33.



authorities collaborate with external stakeholders (e.g. universities) to adapt these tools to their needs (e.g. deploying them to collect data in specific languages). The same approaches could be applied to scale **Claudette** (unfair terms) and **Claudette GDPR** (privacy policies), which can be turned from an investigation tool (e.g. checking a specific clause) to a monitoring tool (e.g. automatically scanning and tracking general terms for unfair clauses, with the methodological reservations mentioned above).

In addition, tools that record and notify when elements change on web pages, such as **Versionista**, can also be used for monitoring purposes. For instance, consumer authorities can manually select hundreds to thousands of websites which they can automatically monitor changes in.

What is more, a lot of the SaaS tools allow user access to APIs (e.g. **DomainTools Iris**, **DomainTools PhishEye**), which entails that further large-scale monitoring can be possible. In this case too, immediate deployment of such options depends on available computer science expertise.

Authorities can also use **AdAnalyst** for monitoring, albeit in a different manner. Some tools (e.g. tools that offer personalized goods, ads, etc.) work on the basis of profiling users. Consumer authorities can use the browser extension developed by the Max Planck Institute for Software Systems to create and train a number of Facebook profiles to mimic different types of consumers. The same can be said for **Ghostery** - the browser extension can be equally employed on the basis of consumer profiles (e.g. selected browsing habits), so consumer authorities can stay up to date with the identities of the companies gathering browsing information from consumers via cookies or other forms of online tracking.

## 4.6 Use the shortlisted investigation and evidence collection tools taking into account their strengths and weaknesses

**VIES** is a reliable, basic tool to check VTA registrations for intracommunity trade, and is already in use by consumer protection authorities, just like **ibancalculator**, a useful tool to obtain the information on the bank that emitted a certain account number.

**Website Evidence Collector** has the advantage that it has been developed by the EDPS, as an EU authority which finds itself directly involved in consumer protection. The tool is open source and can be downloaded and installed, yet computer science knowledge is necessary to deploy it.

**DomainTools Iris** can be used right away for investigating DNS data (e.g. checking registries for domain names, etc.) and the company can further train consumer agencies, although the company's core strengths lay in cybersecurity (e.g. Iris can be used by cybersecurity professionals to detect Remote Access Trojans, or in other words viruses which can be installed on a victim's computer by a cybercriminal in order for the latter to gain access to that later on). This tool can be used together



with the other tool we looked at from DomainTools, namely **PhishEye**, which can search through the world's largest database of malicious domains.

**Scamdoc and Scamadvisor** can be directly used for more manual investigations of scams reported by consumers.

For the authorities with advanced IT knowledge **WHOIS API & Parser System** is the best recommendation possible to give, since the monitoring can be on a high scale.

## 4.7 Invest in the use and development of tools aimed specifically at the needs of consumer authorities with regard to the enforcement of consumer law

When ordering (the development of) custom made tools for the needs of consumer authorities, the following elements are to be taken into account:

- **Task specific but interoperable:** Tools that ought to monitor the compliance with consumer protection rules (e.g. blacklisted unfair commercial practices) need to be developed for specific tasks (e.g. specific practices). However, it may be possible to combine certain tasks. For instance, the same tool could be used to scrape a website and identify whether specific information duties from the Consumer Rights Directive are fulfilled, as well as identify whether a timer is present on the website as an unfair commercial practice. This is possible because the data used for the analysis of the first is the same as for the second (data scraped from a website), and the data analysis is similar (identifying certain code in the website revealing this information, and aggregating observations).
- **The level of harmonisation determines the scope of the tool:** Insofar as fully harmonised rules of consumer law are concerned, a tool with an EU wide scope can be developed. There is the language issue of course, which will need to be taken into account when developing the tool, but the legal assessment remains the same. When the rules of EU consumer law only provide for a minimum level of harmonisation, a tool built to reflect these rules will only catch the most serious infringements, going below the minimum level required by the Directive. National modules will need to be added to reflect the Member States' national rules that offer a higher level of consumer protection.

## 4.8 Accuracy issues with AI tools

One of the most important findings of this report is the difficulty to establish accuracy for most of the tested AI tools. The performance of the tools depends on their intended use and the deployed data. The assessment of computer software is a context-dependent task. Features such as functional completeness, correctness or appropriateness mirror the specified tasks and objectives of the software product, which are in turn determined by the needs of the users. Particular legal issues are

more difficult to identify technically. For instance, in the case of Claudette, accuracy is very difficult to determine, given that general clauses may not be deemed unfair absent (an irrevocable) judicial decision. The higher the legal interpretability, the more likely that accuracy will suffer in such a tool.

Accuracy is also a matter of data quality and human bias, which ought to be taken into account when choosing AI tools that entail training statistical models. The main problem that can arise here is that there is that any bias affecting the data used to train machine learning models will only be replicated by these processes. This issue is increasingly reported in scientific literature, and increasingly being addressed in interdisciplinary, multi-stakeholder conferences such as the ACM Conference on Fairness, Transparency and Accountability (FAccT).<sup>35</sup>

Another issue relating to accuracy deals with evaluation. The less expertise public authorities have with a specific type of technology (e.g. AI tools), the more difficult it becomes for the organization to assess whether claims made by companies in a tendering process accurately reflect the performance of their products and services. Technology companies often misrepresent their solutions and overestimate their accuracy.<sup>36</sup> Such claims can be put into perspective through state of the art academic research. For instance, examples of AI solutions which are far from perfect but gradually improving include 'spam detection, detection of copyright material, hate speech detection, content recommendation'.<sup>37</sup> By comparison, predicting job performance, criminal recidivism or behaviour, or terrorist risk are currently considered fundamentally dubious by scholars from leading computer science institutes.<sup>38</sup> Taking the claims of technology companies for granted, without critically testing and vetting the proposed solution can have serious social and economic implications. However, it must be emphasized that testing sophisticated machine learning tools needs to be done at a much more technical level, directly linked to specific tasks expected to be performed by these tools. For these reasons, more attention needs to be paid to the accuracy debate during tendering processes.

## 4.9 Create European repositories for tools used by authorities

### 4.9.1 Public repositories

**Website Evidence Collector** is a tool that appears on the [EU's Joinup](#) platform (see Figure 6 and Figure 7), together with other tools with some consumer authorities are familiar with, such as the **RAPEX searcher**.<sup>39</sup>

---

<sup>35</sup> <https://facctconference.org/2020/index.html>.

<sup>36</sup> R. Mac et al. (2020), Clearview AI Says Its Facial Recognition Software Identified A Terrorism Suspect. The Cops Say That's Not True, <https://www.buzzfeednews.com/article/ryanmac/clearview-ai-nypd-facial-recognition>.

<sup>37</sup> A. Narayanan, 'How to recognize AI snake oil', <https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf>.

<sup>38</sup> Ibid.

<sup>39</sup> <https://joinup.ec.europa.eu>. As of 7 February 2020, Joinup features 2838 projects.

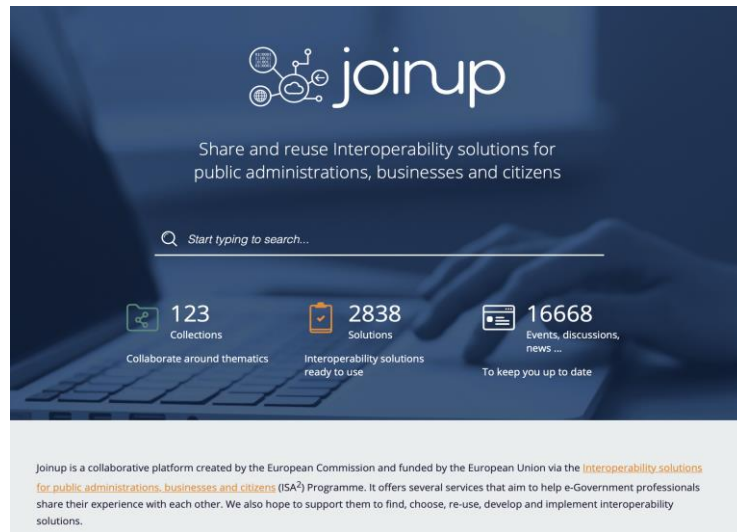


Figure 6: Joinup EU

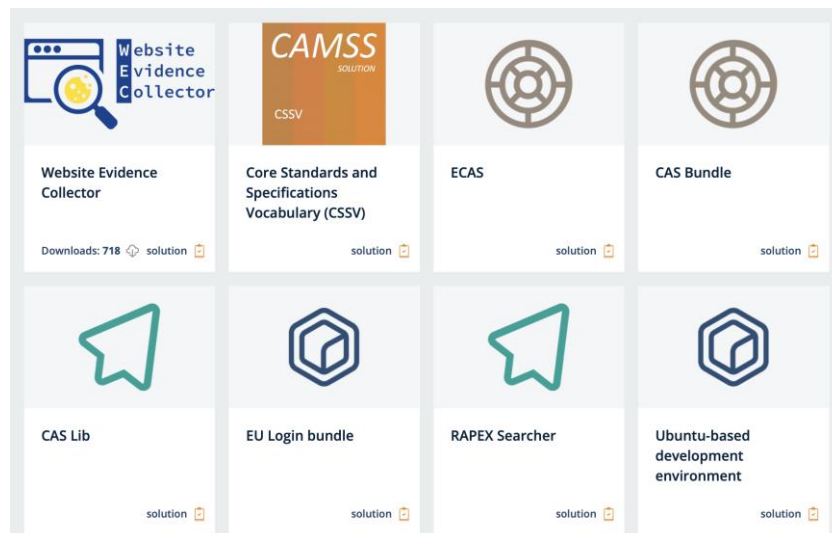


Figure 7: Joinup EU

This platform, however, currently presents three main shortcomings:

- it could benefit from more dissemination among national authorities (not only consumer authorities);
- it features some open source solutions;
- it does not have an optimal structure (e.g. the platform currently allows for browsing, but it does not allow for browsing per category of tool, or per industry, etc.).

In contrast, **Code.gov** (see Figure 8 and Figure 9) is the US federal platform that encourages tool reuse by public administration, and allows for browsing on the basis of department, or also features such as whether the software is open source, etc.<sup>40</sup>

<sup>40</sup> <https://code.gov/>. As of 7 February 2020, Code.gov features 6849 projects.

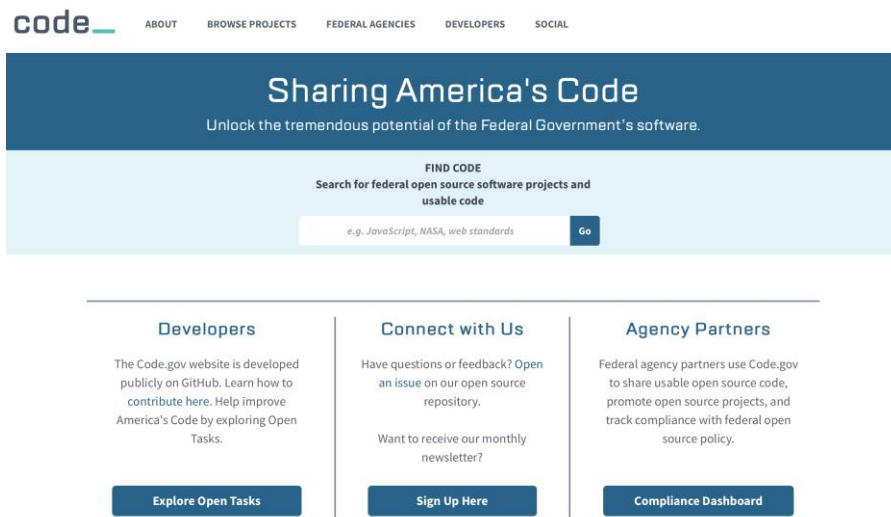


Figure 8: Code.gov

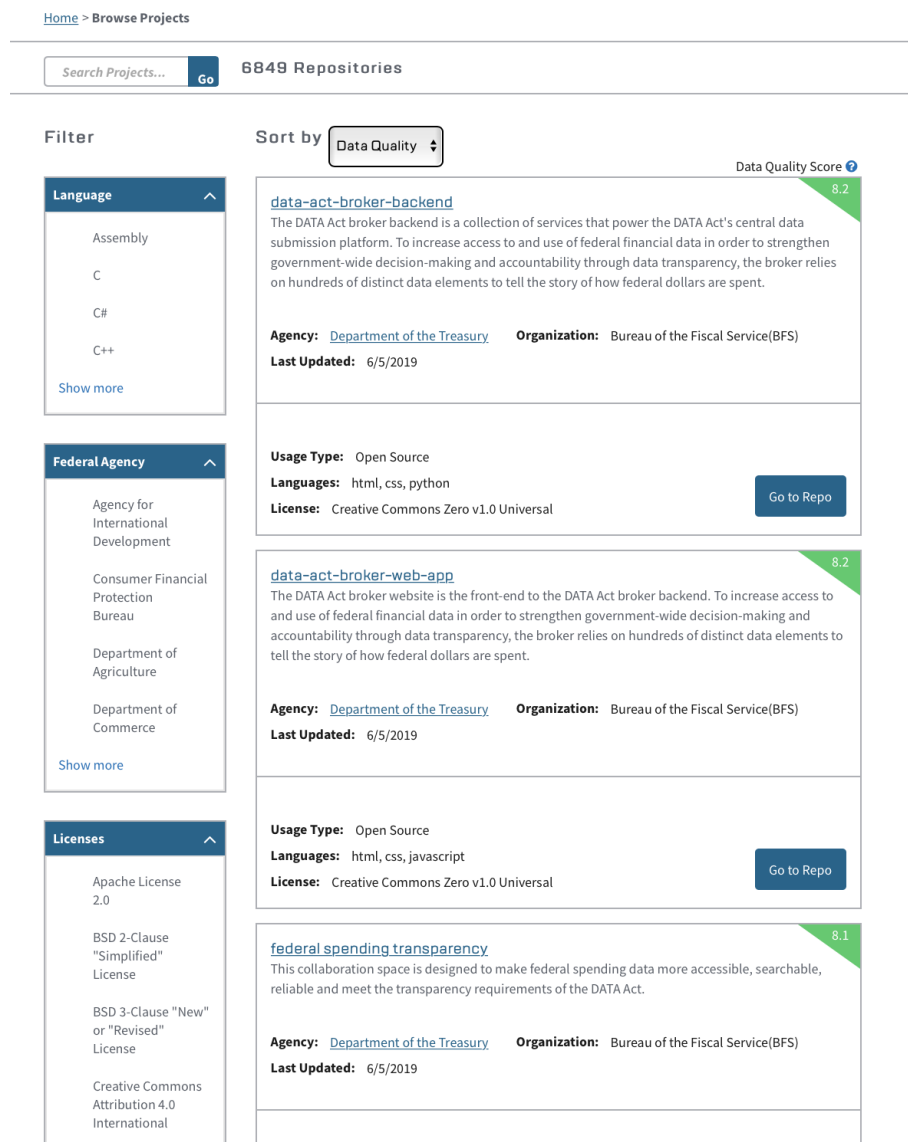


Figure 9: Code.gov

Further developing Joinup as a means of facilitating the sharing and reusing of tools developed nationally or at EU level by consumer authorities should be a priority also in the light of the new coordination framework promoted by the CPC Regulation. We recommend that information about these tools and their intended uses is made findable and accessible (by all EU consumer protection authorities) as advocated by the FAIR (Findable, Accessible, Interoperable and Reusable) principles of data management.<sup>41</sup>

Due consideration must be given to the question of which tools to make available on such a platform. Tools that are fully available (e.g. uploaded on Github, where the full code of the tool is available) can, on the one hand, be fully tested, because of the transparency of the code. However, for critical operations, transparency might also entail the possibility of reverse engineering of tools available publicly, which may lead to those tools becoming redundant or leading to unreliable results.

#### 4.9.2 Internal repositories: interoperable monitoring tools and consumer complaints database infrastructure

This section discusses the implementation of an interoperable consumer complaints storage and retrieval infrastructure for consumer protection authorities.

EU consumer protection agencies currently have disparate methods and ICT systems for allowing consumers to lodge complaints, to investigate them and to resolve them. It is encouraging that the ISA2 initiative (Interoperability solutions for public administrations, businesses and citizens) which supports the Joinup platform mentioned earlier is seeking to facilitate interoperability among these and other infrastructures in public-sector services.<sup>42</sup> Those agencies that still lack mechanisms (databases) for storing complaints digitally, as well as IT tools for categorising, searching and retrieving them, should benefit immensely purely by instating this infrastructure.

The motivation for this would be that such systems would enable software to automatically analyse complaint trends over time, classify complaint types and analyse the frequency of occurrence of certain types. This information can be used to more efficiently allocate resources to address the complaints and streamline complaint resolution processes.

The sheer heterogeneity in the technologies used to build the different AI tools, as well as their licensing terms would make it infeasible to technically integrate them into a single software application. Therefore, our recommendation would be for the commission to develop a user interface ‘dashboard’ that runs in a web browser and instead links users to the chosen tools. However, rather than linking the users to downloads of the tools, we advocate that the tools chosen should be those that are easily adaptable to also run in a web browser (‘in the cloud’). Using the system would

---

<sup>41</sup> Commission Expert Group on Fair Data, Turning FAIR into reality, [http://ec.europa.eu/info/publications/turning-fair-reality\\_en](http://ec.europa.eu/info/publications/turning-fair-reality_en).

<sup>42</sup> [https://ec.europa.eu/isa2/home\\_en](https://ec.europa.eu/isa2/home_en).

then work as follows: the user (consumer protection officer) logs into the consumer rights infringement monitoring dashboard, they are presented with an interface similar to the mock interface depicted below (Figure 10 and Figure 11). Each button on the dashboard links to a chosen AI tool that deals with specific types of infringements. When the user clicks on a particular button, the dashboard logs the user into the chosen AI tools' user interface **which is loaded within the same dashboard screen**.

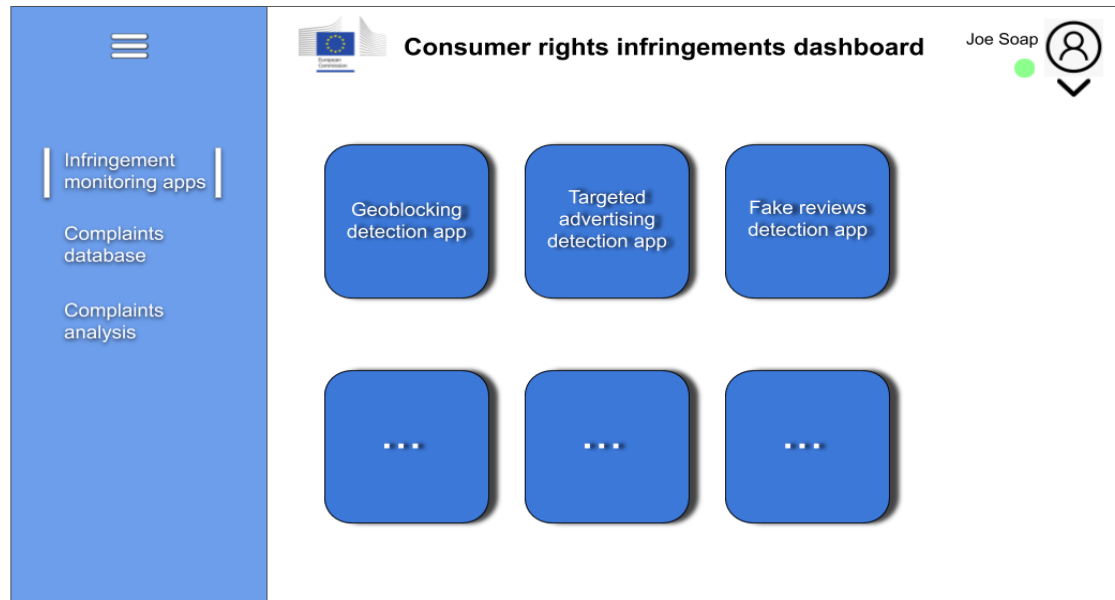


Figure 10: Mock interface to demonstrate an idea for an integrated dashboard of links to consumer rights infringement monitoring tools<sup>43</sup>

<sup>43</sup> A mock interface for an integrated software dashboard of links to selected consumer rights infringement monitoring tools in specific categories.

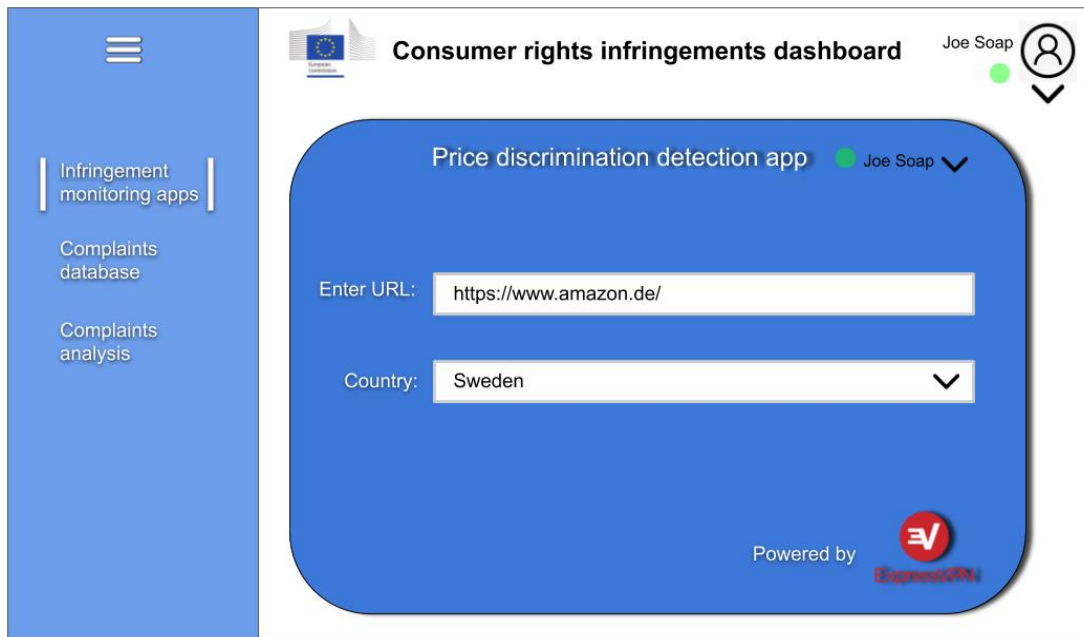


Figure 11: Mock interface of consumer infringement monitoring dashboard that is running a specific selected monitoring tool<sup>44</sup>

This would require less software modifications required by the different AI tools. If the commission partners with the developers of the chosen tools, they could also request them to all build a common, uniform, user interface for their tools that is compatible (from a look and feel perspective) with the consumer rights infringement monitoring dashboard. This is a more feasible approach because it would require much less development hours and coordination than integrating all the tools into a single software application.

#### 4.9.3 Connect the monitoring tools dashboard with the consumer complaints database of each relevant EU consumer protection authority

The tools recommended in our study have a data analysis component to them to monitor or detect behaviour by eCommerce and web service platforms that could lead to consumer rights infringements. If the information and insights generated by these tools could be automatically shared with, or imported into, consumer complaints databases, it would enable automated comparison of detected infringements with received complaints. This could, for example, allow identification of those offenders and offenses that are not being detected by consumers (i.e. the more subtle infringements). This could, in turn, inform strategies for consumer rights awareness campaigns.

<sup>44</sup> Mock interface of a fictional consumer infringement monitoring dashboard that could run ExpressVPN (with a tailored user interface) as its chosen price discrimination detection app.



## 5 List of literature

Allen, M. and Cervo, D., 'Data Quality Management' in Multi-Domain Master Data Management Advanced MDM and Data Governance in Practice 2015, p. 131-160;

Cisco, 'What Is Software as a Service', available at <https://www.cisco.com/c/en/us/products/software/what-is-software-as-a-service-saas.html>;

Contissa et al. (2018) Towards Consumer-Empowering Artificial Intelligence, International Joint Conference on Artificial Intelligence (IJCAI) <http://www.ijcai.org>.

Commission Expert Group on Fair Data, Turning FAIR into reality, [http://ec.europa.eu/info/publications/turning-fair-reality\\_en](http://ec.europa.eu/info/publications/turning-fair-reality_en).

Ducato, R. and Strowel A.M., (2018) Limitations to Text and Data Mining and Consumer Empowerment: Making the Case for a Right to Machine Legibility, [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3278901](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3278901);

High-Level Expert Group on AI, Ethics Guidelines for Trustworthy Artificial Intelligence, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>.

Hunt, T., 'Data Enrichment, People Data Labs and Another 622M Email Addresses', <https://www.troyhunt.com/data-enrichment-people-data-labs-and-another-622m-email-addresses/>.

IDG, 'Cloud computing survey' (IDG Communications, 2018), available at <https://cdn2.hubspot.net/hubfs/1624046/2018%20Cloud%20Computing%20Executive%20Summary.pdf>.

Lippi, M., Contissa, G., Lagioia, F. et al. (2019) Consumer protection requires artificial intelligence. *Nat Mach Intell* 1, 168–169 doi:10.1038/s42256-019-0042-3.

Lippi, M., Pałka, P., Contissa, G. et al. (2019) CLAUDETTE: an automated detector of potentially unfair clauses in online terms of service. *Artif Intell Law* 27: 117.

Ibrahamin, O. and Larsson, A. (2016) 'A systems tool for structuring public policy problems and design of policy options', *Int. J. Electronic Governance*, Vol. 9, Nos. 1/2, 4-26.

Ibrahamin, O. and Larsson, A. (2015) Policy Modelling and Simulation Tool, [http://www.sense4us.eu/images/reports/611242-Sense4us\\_D6-2-Policy\\_Modelling\\_Tool-FINAL.pdf](http://www.sense4us.eu/images/reports/611242-Sense4us_D6-2-Policy_Modelling_Tool-FINAL.pdf).

Jansen, M. and Joha, A. (2012), 'Adoption Challenges of Introducing and Implementing SaaS', in Mahmud Akhter Shareef, Norm Archer, Yogesh K. Dwivedi, *Transformational Government Through EGov Practice: Socio-Economic, Cultural, and Technological Issues* (Emerald Publishing).



Mac, R., et al. (2020), Clearview AI Says Its Facial Recognition Software Identified A Terrorism Suspect. The Cops Say That's Not True, <https://www.buzzfeednews.com/article/ryanmac/clearview-ai-nypd-facial-recognition>.

Mathur, Arunesh et al. (2019) Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites. Proc. ACM Hum.-Comput. Interact. 3, CSCW, Article 81 (November 2019), 32 pages. DOI: <https://doi.org/10.1145/3359183>.

Micklitz et al. (2017) The Empire Strikes Back: Digital Control of Unfair Terms of Online Services, Journal of Consumer Policy, vol. 40, issue 3, 367-388.

Narayanan, A., 'How to recognize AI snake oil', <https://www.cs.princeton.edu/~arvindn/talks/MIT-STS-AI-snakeoil.pdf>.

Norris, D.F. et al., 'Cyberattacks at the Grass Roots: American Local Governments and the Need for High Levels of Cybersecurity' (2019) 79(6) Public Administration Review 895.

Sylos, M., 'Top five advantages of software as a service (SaaS)' (IBM, 18 September 2013), <https://www.ibm.com/blogs/cloud-computing/2013/09/18/top-five-advantages-of-software-as-a-service-saas/>.

von der Leyen, U., Political guidelines for the next European Commission 2019-2024, [https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission\\_en.pdf](https://ec.europa.eu/commission/sites/beta-political/files/political-guidelines-next-commission_en.pdf).

## 6 List of Annexes

Annex 1	In-depth descriptions of 25 tools and narratives
Annex 2	Questionnaire for consumer agencies/organizations
Annex 3	Questionnaire for researchers
Annex 4	List of consumer agencies/organizations contacted
Annex 5	List of search strings and categories
Annex 6	Number of <i>potentially</i> relevant tools depending on the source
Annex 7	Overview and a list of 73 relevant tools

## Annex 2 Questionnaire for consumer agencies/organizations

### Welcome screen

On behalf of the European Commission, Maastricht University (The Maastricht Law and Tech Lab and the Institute of Data Science) currently conducts research on AI/IT tools used by consumer authorities:

- for the purpose of digital market surveillance in the EU, e.g. discovering unfair terms or unfair commercial practices in digital markets, fake reviews, personalised pricing, non-compliance with information duties, illegal geoblocking etc.,
- in order to find out about practices that are not regulated yet, but are a cause of consumer concerns which they express online.

We would like to ask you a couple of questions about these tools!

Please note that your responses will be used solely for the purpose of this research. They will be kept confidential, stored securely, and made available only to the members of the research team. The responses will be anonymized for the purpose of preparing the final report for the European Commission. Both - the European Commission and Maastricht University - agreed to keep all information or documents related to the performance of this research confidential.

### Consent

- By clicking this box I agree that my first and last name as well as my email address is used to distribute this survey and collect answers to it. Please see our Privacy Notice for more details on how we handle your data. (5)

Q1 Does your organization/institution monitor online markets or carry out internet investigations to detect infringements of consumer protection laws?

- Yes (1)
- No (2)
- I do not know (3)

*Display This Question:*

*If Does your organization/institution monitor online markets or carry out internet investigations to... = No*

Q2 Why, in your opinion, such activities are not carried out?

---

---

---

---

---

*Display This Question:*

*If Does your organization/institution monitor online markets or carry out internet investigations to... = No*

Q3 Are there, to your knowledge, any plans to undertake such activities in the future?

- Definitely yes (1)
- Probably yes (2)
- Might or might not (3)
- Probably not (4)
- Definitely not (5)

*Display This Question:*

*If Does your organization/institution monitor online markets or carry out internet investigations to... = Yes*

Q4 Does your institution use any digital tools to carry out these activities?

- Yes (1)
- No (2)

*Display This Question:*

*If Does your institution use any digital tools to carry out these activities? = No*

Q5

We would still like to learn a few more things about the needs of your institution/organization!

---

*Display This Question:*

*If Does your institution use any digital tools to carry out these activities? = No*

Q6 How does your institution/organization undertake investigations related to harms affecting online consumers?

---

---

---

---

---

---

*Display This Question:*

*If Does your institution use any digital tools to carry out these activities? = No*

Q7

What are the most prominent issues your institution/organization faces when undertaking online investigations?

- 1 (1) \_\_\_\_\_
- 2 (2) \_\_\_\_\_
- 3 (3) \_\_\_\_\_
- 4 (4) \_\_\_\_\_
- 5 (5) \_\_\_\_\_

*Display This Question:*

*If Does your institution use any digital tools to carry out these activities? = Yes*

Q8 Please list all tools that are used by your organization/institution!

	Name of a tool	How can we find it? (website url, reference to a research paper etc.)
	(1)	(1)
1 (1)		
2 (2)		
3 (3)		
4 (4)		
5 (5)		

*Display This Question:*

*If Does your institution use any digital tools to carry out these activities? = Yes*

Q20 Are you aware of any other tools that could be used to carry out such activities?

	Name of a tool	How can we find it? (website url, reference to a research paper etc.)	Please check the box if your organization/institution plans on using this tool in the future.
	(1)	(1)	(1)
1 (1)			•
2 (2)			•

3 (3)			•
4 (4)			•
5 (5)			•

*Display This Question:*

*If Does your institution use any digital tools to carry out these activities? = Yes*

Q21 What criteria do you think are relevant for the quality assessment of these tools?

- 1 (1) \_\_\_\_\_
- 2 (2) \_\_\_\_\_
- 3 (3) \_\_\_\_\_
- 4 (4) \_\_\_\_\_
- 5 (5) \_\_\_\_\_

*Display This Question:*

*If Does your organization/institution monitor online markets or carry out internet investigations to... = No*

Q17 Are you aware of any AI or IT tools that could be used to carry out such activities?

Please list all tools that you have heard about.

	Name of a tool	How can we find it? (website url, reference to a research paper etc.)
	(1)	(1)
1 (1)		
2 (2)		
3 (3)		
4 (4)		
5 (5)		

Q9 Are you aware of any other organizations/institutions in your country that carry out online market monitoring that may be relevant for detecting consumer law infringements?

- Yes (1)
- No (2)

*Display This Question:*

*If Are you aware of any other organizations/institutions in your country that carry out online marke... = Yes*

Q10 Please list all these organizations/institutions

- 1 (1) \_\_\_\_\_
- 2 (2) \_\_\_\_\_
- 3 (3) \_\_\_\_\_
- 4 (4) \_\_\_\_\_
- 5 (5) \_\_\_\_\_

Q11 Could you point us to a person who could provide us with more information on such investigations/tools or institutions using them?

- Yes (1)
- No (2)

---

*Display This Question:*

*If Could you point us to a person who could provide us with more information on such investigations/... = Yes*

Q12 Please give us contact details.

	Name of a person	Contact details (institution, e-mail, phone number)
	(1)	(1)
1 (1)		
2 (2)		
3 (3)		
4 (4)		
5 (5)		

Q18 Do you want to share anything else that you think could be relevant for our research?

---

---

---

---

---

---

*Display This Question:*

*If Does your institution use any digital tools to carry out these activities? = Yes*

Q13 Would you be available for a phone/Skype interview to tell us more about the activity of your organization/institution in dealing with online consumer harms?

- Yes (1)
- No (2)

---

*Display This Question:*

*If Would you be available for a phone/Skype interview to tell us more about the activity of your org... = Yes*

Q14 Could you provide us with your contact details!

- First name (1) \_\_\_\_\_
- Last name (2) \_\_\_\_\_
- Email address (3) \_\_\_\_\_



## Annex 3 Questionnaire for researchers

### Welcome screen

Thank you for taking the time to fill out this short survey!

On behalf of the European Commission, Maastricht University (The Maastricht Law and Tech Lab and the Institute of Data Science) currently conducts research on the digital enforcement of European consumer protection law.

The purpose of our study is to learn about existing information technology (IT) / artificial intelligence (AI) tools that are or can be made useful in online market surveillance for consumer policy development purposes and for the enforcement of consumer protection legislation.

### We would like to ask you what such tools are known to you!

In the following, you will see a form that will allow you to fill out the names of and references to the tools that you know. Please interpret the above description of 'tools' very broadly! List anything that comes to your mind and you think might be relevant for us!

Please note that your responses will be used solely for the purpose of this research. They will be kept confidential, stored securely, and made available only to the members of the research team. The responses will be anonymized for the purpose of preparing the final report for the European Commission.

- By clicking this box I agree that my first and last name as well as my email address is used to distribute this survey and collect answers to it.  
Please see our Privacy Notice for more details on how we handle your data.

Q1 Please list any tool that you think we might want to learn about!

	Name of a tool	How can we find it? (website url, reference to a research paper etc.)
1		
2		
3		
4		
5		

Q2 What criteria do you think are relevant for the quality assessment of these tools?

- 1. \_\_\_\_\_
- 2. \_\_\_\_\_
- 3. \_\_\_\_\_
- 4. \_\_\_\_\_
- 5. \_\_\_\_\_

Q3 Do you know any other resources that could be relevant for us? Do you want to share anything else that you think could be relevant for our research?

---

---

---

---

---

---

Q4 Do you want to be updated about this research?

- Yes
- No

Q5 Please fill out your contact details (*optional*):

- First name \_\_\_\_\_
- Last name \_\_\_\_\_
- Email address \_\_\_\_\_

## **Annex 4 List of consumer agencies/organizations contacted**

1. BEUC - The European Consumer Organization
2. DGCCRF - Direction générale de la concurrence, de la consommation et de la répression des fraudes (France)
3. National Trading Standards (UK)
4. Finnish Competition and Consumer Authority (Finland)
5. The Consumer Protection and Technical Regulatory Authority (Estonia)
6. The Norwegian Consumer Authority (Norway)
7. Swedish Consumer Agency - Konsumentverket (Sweden)
8. The Authority for Consumers and Markets (Netherlands)
9. Verbraucherzentrale Bundesverband (vzbv) (the Federation of German Consumer Organisations)
10. Federal Ministry of Justice and Consumer Protection (Germany)
11. International Consumer Protection Enforcement Network (ICPEN)
12. Better Business Bureau
13. Electronic Frontier Foundation
14. W3C
15. KU Leuven Center for IT & IP Law
16. COSIC
17. None of Your Business
18. Stichting Reclame Code
19. IPFS
20. Coala global

## Annex 5 List of search strings and categories

Category	Description and definitions	Topic	keywords
Web Evidence	tools that help to record different states of the internet	Internet archives	wayback machine
			internet snap
			internet versions
			internet archive
		Webpage snaps	webpage snap
			programmed screenshot
			automatic screenshot
			screenshot
		Digital investigations	digital investigations
		Domain registries	domain validity
domain registries			
domain monitoring			
Advertising	practices used by traders (either directly or through paid intermediaries) in order to increase the sale of their products or services	Online trackers	market monitoring
		Behavioral tracking	consumer profiling algorithm
			behavioral profiling
			behavioral targeting
			targeted advertising
			website monitoring
			website tracking
		HTTP Cookies	cookie monitoring
			cookie tracking
		Misleading advertising	ad blocking
			ad profiling
			false advertising
illegal advertising			
misleading advertising			

			sponsored advertising
		Social media issues	sponsored advertising
			fake followers
			influencer disclosure
			influencer marketing
		Fake reviews	detecting fake reviews
			fake review detection
			tracking fake reviews
		Web beacon tracker	web beacon behavioral tracking
			web beacon
<b>Dark Patterns</b>	user interface design choices that benefit an online service by coercing, steering, or deceiving users into making unintended and potentially harmful decisions	Sneaking	hidden cost detection
			hidden subscription block
			sneak into basket block
		Urgency	countdown timer
			limited time message
		Misdirection	confirm shaming
			pressured selling
		Social proof	fake testimonials
			fake testimonial detection
		Scarcity	low stock message
			scarcity cues
			detect high demand message
		Obstruction	hard to cancel subscription
		Forced Action	forced enrollment websites
			forced enrollment detection

			website timer
<b>Geoblocking</b>	is technology that restricts access to Internet content based upon the user's geographical location	Geoblocking	geoblocking
			ip blocking
			privacy browser
		VPN	privacy extension
			vpn alternative
			free vpn
			change ip vpn
<b>Scam (Online Fraud)</b>	other (illegal) methods used to unfairly disadvantage consumers	Financial online fraud	cloned firms scam
			boiler room scam
			recovery room scam
			credit scam
			auction scam
			fake credit
		Identity Theft	fake consumer
			fake consumer detection
			Identity theft
			Identity theft detection
			Identity theft monitor
			identity protection
			monitoring identity theft
		General scam & Phishing	phishing
			phishing detection
			phishing monitoring
			detecting e-commerce
			e-commerce fraud
			online law enforcement
			online shopping fraud
			online shopping law enforcement
			scam
			scam monitoring
			scam detection

			scam monitor
			detect fake advise
			detect fake advise
			ai consumer fraud online
<b>Unfair terms issues</b>	contract terms that significantly limit the rights of the consumer under the applicable law	Consumer protection to unfair terms	analyzing privacy policies
			consumer protection
		General terms and conditions	terms of service
			unfair terms
<b>Price Transparency Issues</b>	sing different prices for different types of customers without making clear the criteria that determine the difference	Personal pricing	individual pricing
			monitoring individual pricing
			monitoring personal pricing
			personal pricing
			e-commerce fraud pricing
		Price discrimination	algorithm detect unfair pricing website
			price discrimination
			price discrimination detection
			product price change
			price changes detection
			price discrimination algorithm
<b>Compliance</b>	all issues related with non-compliance of online services, and not taking reasonable measures to allow impaired persons to use the website	Withdrawal issues	short withdrawal period
			withdrawal rights
			withdrawal rights disclaimer
		Failure of Web accessibility	Internet accessibility
			web accessibility
		Compliance and obligations	mandatory disclosures detection

			non-compliance with information obligations
			do-not-call registry



## Annex 6 Number of potentially relevant tools depending on the source

Sources	Total number of tools extracted	Number of filtered tools before the assessment	Number of tools assessed and recommended
Questionnaires	64	19	9
Database search	355 (sampled from 5030)	28	2
Desktop search	56	14	6
Word of mouth	42	12	8
<b>Total</b>	<b>517</b>	<b>73</b>	<b>25</b>

## Annex 7 A list of 73 relevant tools

	Source	Category
1.	<a href="https://worldwide.espacenet.com/publicationDetails/biblio?CC=US&amp;NR=6950804&amp;KC=B2&amp;date=20050927&amp;locale=en_EP">https://worldwide.espacenet.com/publicationDetails/biblio?CC=US&amp;NR=6950804&amp;KC=B2&amp;date=20050927&amp;locale=en_EP</a>	Advertising
2.	<a href="http://arxiv.org/abs/1407.0697v1">http://arxiv.org/abs/1407.0697v1</a>	Advertising
3.	<a href="https://www.domaintools.com/products/iris/">https://www.domaintools.com/products/iris/</a>	Compliance
4.	<a href="http://vendors.r2accelerator.org">http://vendors.r2accelerator.org</a>	Compliance
5.	<a href="https://www.watchlist-internet.at/">https://www.watchlist-internet.at/</a>	Consumer complaints
6.	<a href="https://github.com/aruneshmathur/dark-patterns">https://github.com/aruneshmathur/dark-patterns</a>	Dark Patterns
7.	<a href="https://darkpatterns.uxp2.com">https://darkpatterns.uxp2.com</a>	Dark Patterns
8.	<a href="http://arxiv.org/abs/1109.1074v1">http://arxiv.org/abs/1109.1074v1</a>	Dark Patterns
9.	<a href="https://www.spyfu.com/?alt=1">https://www.spyfu.com/?alt=1</a>	Web Evidence
10.	<a href="https://distill.io">https://distill.io</a>	Web Evidence
11.	<a href="http://whois.domaintools.com">http://whois.domaintools.com</a>	Web Evidence
12.	<a href="https://reviewmeta.com">https://reviewmeta.com</a>	Advertising
13.	<a href="http://www.bodacc.fr/">http://www.bodacc.fr/</a>	Web Evidence
14.	<a href="https://www.infogreffe.fr/recherche-siret-entreprise/chercher-siret-entreprise.html">https://www.infogreffe.fr/recherche-siret-entreprise/chercher-siret-entreprise.html</a>	Web Evidence
15.	<a href="http://www.iana.org/numbers">www.iana.org/numbers</a>	Web Evidence
16.	<a href="https://dnslytics.com">dnslytics.com</a>	Web Evidence
17.	<a href="https://www.fakespot.com/">https://www.fakespot.com/</a>	Advertising
18.	<a href="https://www.domaintools.com/products/phisheye/">https://www.domaintools.com/products/phisheye/</a>	Scam (Online Fraud)
19.	<a href="http://mxtoolbox.com">mxtoolbox.com</a>	Web Evidence
20.	<a href="http://arxiv.org/abs/1904.12607">http://arxiv.org/abs/1904.12607</a>	Advertising

21.	<a href="https://dl.acm.org/citation.cfm?id=2783370">https://dl.acm.org/citation.cfm?id=2783370</a>	Advertising
22.	<a href="https://dl.acm.org/citation.cfm?id=2070716">https://dl.acm.org/citation.cfm?id=2070716</a>	Advertising
23.	<a href="http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=290&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=1&amp;cacheurlFromRightClick=no">http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=290&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=1&amp;cacheurlFromRightClick=no</a>	Advertising
24.	<a href="https://www.scamadviser.com">https://www.scamadviser.com</a>	Scam (Online Fraud)
25.	<a href="http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=290&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=10&amp;cacheurlFromRightClick=no">http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=290&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=10&amp;cacheurlFromRightClick=no</a>	Advertising
26.	<a href="http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=290&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=2&amp;doc=11&amp;cacheurlFromRightClick=no">http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=290&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=2&amp;doc=11&amp;cacheurlFromRightClick=no</a>	Advertising
27.	<a href="http://claudette.eui.eu/">http://claudette.eui.eu/</a>	Unfair terms issues
28.	<a href="http://arxiv.org/abs/1607.06891v3">http://arxiv.org/abs/1607.06891v3</a>	Scam (Online Fraud)
29.	<a href="https://github.com/DocNow/waybackproof">https://github.com/DocNow/waybackproof</a>	Web Evidence
30.	<a href="https://github.com/rodcoelho/dotfiles">https://github.com/rodcoelho/dotfiles</a>	Price Transparency Issues
31.	<a href="https://www.expressvpn.com">https://www.expressvpn.com</a>	Geoblocking
32.	<a href="https://doi.org/10.1016/j.dss.2016.04.003">https://doi.org/10.1016/j.dss.2016.04.003</a>	Scam (Online Fraud)
33.	<a href="https://arxiv.org/abs/1808.07293">https://arxiv.org/abs/1808.07293</a>	Advertising
34.	<a href="http://arxiv.org/abs/1907.03048v1">http://arxiv.org/abs/1907.03048v1</a>	Scam (Online Fraud)
35.	<a href="https://play.google.com/store/apps/details?id=com.vpn.bestSharkVPN&amp;referrer=utm_source%3D42matters.com%26utm_medium%3Dapi">https://play.google.com/store/apps/details?id=com.vpn.bestSharkVPN&amp;referrer=utm_source%3D42matters.com%26utm_medium%3Dapi</a>	Scam (Online Fraud)
36.	<a href="https://doi.org/10.1016/j.elerap.2013.01.001">https://doi.org/10.1016/j.elerap.2013.01.001</a>	Scam (Online Fraud)

37.	<a href="https://worldwide.espacenet.com/publicationDetails/biblio?CC=US&amp;NR=8565396&amp;KC=B1&amp;date=20131022&amp;locale=en_EP">https://worldwide.espacenet.com/publicationDetails/biblio?CC=US&amp;NR=8565396&amp;KC=B1&amp;date=20131022&amp;locale=en_EP</a>	Scam (Online Fraud)
38.	<a href="https://github.com/juliankrieger/Screengrabber">https://github.com/juliankrieger/Screengrabber</a>	Web Evidence
39.	<a href="https://visualping.io/cdlp.html?utm_source=cd&amp;utm_medium=redirtocdlp&amp;utm_campaign=mig1/">https://visualping.io/cdlp.html?utm_source=cd&amp;utm_medium=redirtocdlp&amp;utm_campaign=mig1/</a>	Web Evidence
40.	<a href="http://versionista.com">versionista.com</a>	Web Evidence
41.	<a href="https://www.domaintools.com/products/domain-risk-score/">https://www.domaintools.com/products/domain-risk-score/</a>	Web Evidence
42.	<a href="https://tosdr.org/about.html">https://tosdr.org/about.html</a>	Unfair terms issues
43.	<a href="http://www.itiis.org/digital-library/manuscript/355">http://www.itiis.org/digital-library/manuscript/355</a>	Geoblocking
44.	<a href="https://censoredplanet.org/assets/403forbidden.pdf">https://censoredplanet.org/assets/403forbidden.pdf</a>	Geoblocking
45.	<a href="https://drs.whoisxmlapi.com/domain-monitor">https://drs.whoisxmlapi.com/domain-monitor</a>	Web Evidence
46.	<a href="https://chrome.google.com/webstore/detail/reveye-reverse-image-search/keaaclicjhehbbapnphnmpiklalfhelgf/">https://chrome.google.com/webstore/detail/reveye-reverse-image-search/keaaclicjhehbbapnphnmpiklalfhelgf/</a>	Web Evidence
47.	<a href="https://www.whois.net/default.aspx">https://www.whois.net/default.aspx</a>	Web Evidence
48.	<a href="https://www.ibancalculator.com/">https://www.ibancalculator.com/</a>	Scam (Online Fraud)
49.	<a href="https://play.google.com/store/apps/details?id=se.christofferalf.webpagemonitor&amp;referrer=utm_source%3D42matters.com%26utm_medium%3Dapi">https://play.google.com/store/apps/details?id=se.christofferalf.webpagemonitor&amp;referrer=utm_source%3D42matters.com%26utm_medium%3Dapi</a>	Advertising
50.	<a href="http://www.bindb.com/bin-database.html">www.bindb.com/bin-database.html</a>	Scam (Online Fraud)
51.	<a href="http://ec.europa.eu/taxation_customs/vies/vieshome.do?locale=en">http://ec.europa.eu/taxation_customs/vies/vieshome.do?locale=en</a>	Scam (Online Fraud)
52.	<a href="http://arxiv.org/abs/1806.08910v1">http://arxiv.org/abs/1806.08910v1</a>	Scam (Online Fraud)
53.	<a href="https://dl.acm.org/citation.cfm?id=1529575">https://dl.acm.org/citation.cfm?id=1529575</a>	Price Transparency Issues

54.	<a href="https://www.econstor.eu/handle/10419/181294">https://www.econstor.eu/handle/10419/181294</a>	Price Transparency Issues
55.	<a href="http://arxiv.org/abs/1711.06955v1">http://arxiv.org/abs/1711.06955v1</a>	Price Transparency Issues
56.	<a href="http://arxiv.org/abs/1903.11469">http://arxiv.org/abs/1903.11469</a>	Price Transparency Issues
57.	<a href="http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=244&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=10&amp;cacheurlFromRightClick=no">http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=244&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=10&amp;cacheurlFromRightClick=no</a>	Scam (Online Fraud)
58.	<a href="https://www.ftc.gov/system/files/documents/public_comments/2015/09/00011-97593.pdf">https://www.ftc.gov/system/files/documents/public_comments/2015/09/00011-97593.pdf</a>	Price Transparency Issues
59.	<a href="https://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=4&amp;SID=8DL6HUjKqtC2WF3qPtf&amp;page=1&amp;doc=3&amp;cacheurlFromRightClick=no">https://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=4&amp;SID=8DL6HUjKqtC2WF3qPtf&amp;page=1&amp;doc=3&amp;cacheurlFromRightClick=no</a>	Advertising
60.	<a href="https://nordvpn.com">https://nordvpn.com</a>	Geoblocking
61.	<a href="http://arxiv.org/abs/1605.05077">http://arxiv.org/abs/1605.05077</a>	Advertising
62.	<a href="https://www.f-secure.com/en/home/products/freedom">https://www.f-secure.com/en/home/products/freedom</a>	Geoblocking
63.	<a href="https://play.google.com/store/apps/details?id=com.usefullapps.fakegpslocationpro&amp;hl=en">https://play.google.com/store/apps/details?id=com.usefullapps.fakegpslocationpro&amp;hl=en</a>	Advertising
64.	<a href="https://www.ghostery.com/wp-content/themes/ghostery/images/campaigns/tracker-study/Ghostery_Study_-_Tracking_the_Trackers.pdf">https://www.ghostery.com/wp-content/themes/ghostery/images/campaigns/tracker-study/Ghostery_Study_-_Tracking_the_Trackers.pdf</a>	Advertising
65.	<a href="http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=239&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=2&amp;cacheurlFromRightClick=no">http://apps.webofknowledge.com/full_record.do?product=WOS&amp;search_mode=GeneralSearch&amp;qid=239&amp;SID=6BN E5W34w9zUGQOFJug&amp;page=1&amp;doc=2&amp;cacheurlFromRightClick=no</a>	Advertising
66.	<a href="https://github.com/lassebunk/screenshot-generator">https://github.com/lassebunk/screenshot-generator</a>	Web Evidence
67.	<a href="https://github.com/mozilla/OpenWPM">https://github.com/mozilla/OpenWPM</a>	Advertising

68.	<a href="https://github.com/BBC-News/wraith">https://github.com/BBC-News/wraith</a>	Web Evidence
69.	<a href="https://www.aignes.com">https://www.aignes.com</a>	Web Evidence
70.	<a href="https://ghostery.com">ghostery.com</a>	Advertising
71.	<a href="https://www.techsmith.com">https://www.techsmith.com</a>	Web Evidence
72.	<a href="https://play.google.com/store/apps/details?id=com.codefy.beacons&amp;referrer=utm_source%3D42matters.com%26utm_medium%3Dapi">https://play.google.com/store/apps/details?id=com.codefy.beacons&amp;referrer=utm_source%3D42matters.com%26utm_medium%3Dapi</a>	Advertising
73.	<a href="https://osirtbrowser.com">osirtbrowser.com</a>	Web Evidence
74.	Claudette GDPR	
75.	<a href="https://github.com/LCS2-IIITD/DeFrauder">https://github.com/LCS2-IIITD/DeFrauder</a>	
76.	<a href="https://www.wappalyzer.com/">https://www.wappalyzer.com/</a>	
77.	scamdoc	
78.	<a href="https://chrome.google.com/webstore/detail/wot-web-of-trust-website/bhmmomiinigofkjcapedijndpbi kbInp">https://chrome.google.com/webstore/detail/wot-web-of-trust-website/bhmmomiinigofkjcapedijndpbi kbInp</a>	
79.	<a href="https://www.uppsalasecurity.com/index.html">https://www.uppsalasecurity.com/index.html</a>	
80.	Affiliate Marketing Disclosures	Advertising
81.	<a href="https://wellkeptwallet.com/get-paid-to-write-reviews/">https://wellkeptwallet.com/get-paid-to-write-reviews/</a>	
82.	EDPS	
83.	Sentinel	
84.	EDPS Website Evidence Collector	Web evidence
85.	INRIA cookie tool	Web evidence