Computer Law & Security Review 56 (2025) 106080

Contents lists available at ScienceDirect

Computer Law & Security Review: The International Journal of Technology Law and Practice

journal homepage: www.elsevier.com/locate/clsr

Botnet defense under EU data protection law

Piotr Rataj *

Zentrum für Recht und Digitalisierung (ZRD) Saar, Saarbrücken, Saarland Informatics Campus, Saarbrücken, Germany

ARTICLE INFO

Keywords:
Botnet defense
data protection law
cybersecurity regulation
information sharing
IP address

ABSTRACT

We analyse the legal framework spanned by EU data protection law with respect to the defence against botnet-related threats. In particular, we examine what legal constraints the General Data Protection Regulation (GDPR) (and others) impose on the processing of personal data when that processing aims at detecting botnet-related traffic. We thereby put data protection rules into perspective with current trends in European IT security regulation, specifically Directive 2022/2555/EU (NIS 2 Directive).

We find that the resulting legal landscape is complex and has not yet been sufficiently explored. Our analysis provides an initial evaluation of a wide range of emerging legal issues. In particular, we consider four typical processing scenarios, such as DNS sinkholing by a public authority or sharing of cybersecurity-related personal data, and discuss some of their legal problems, linking them as thoroughly as possible to potentially relevant case law of the European Court of Justice.

1. Introduction

Botnet defence is a major concern for network security, given that a constantly growing number of computing devices are connected to the Internet and thus can be subject to attacks turning them into bots. A botnet is essentially a set of computing devices infected by malware (bots) whose computations are orchestrated remotely by a malicious source (botmaster). While these computations are typically not obvious to the users of the infected machines, the computational resources thus acquired are then often used by the controller to conduct (further) cyber-attacks. These attacks may be directed against the infected machine itself or other machines (over the Internet). A paradigmatic example of the latter is a distributed denial-of-service (DDoS) attack, in which multiple bots are triggered to simultaneously flood a victim machine with requests, eventually exhausting the machine's connectivity resources.

From the perspective of EU law, Member States must already criminalise the act of assembling a botnet.[1] The obligation to introduce criminal sanctions also applies with regard to cyberattacks typically mediated by a botnet, such as the aforementioned DDoS attacks. However, due to the distributed structure of botnets, intervention against attackers by means of law enforcement faces serious obstacles in practice. In particular, botnets typically span across multiple countries, so that different jurisdictions and their institutions must get involved. Usually, this requires cumbersome procedures.[2] In contrast, botnets can adapt quickly (for instance, by restructuring the command channels) and thus jeopardize legal enforcement efforts. In addition, botmasters typically use elaborated techniques to obfuscate their identities and locations (e.g., the Tor network),[3] which makes tracking them even harder. Consequently, the number of successful prosecutions is

* Corresponding author.
E-mail address: piotr.rataj@zrd-saar.de.

[1] Article 3, read in conjunction with Recital 5, of the Directive 2013/40/EU of the European Parliament and of the Council of 12 August 2013 on attacks against information systems and replacing Council Framework Decision 2005/222/JHA, [2013] OJ L218/8.

[2] G. Gerard, 'Botnet Mitigation and International Law' (2019) 58 Columbia Journal of Transnational Law 189, 200-201, therefore calling for fixes in international law. A recent step to improve on this problem can be seen in the Council of Europe's Second Additional Protocol to the Cybercrime Convention on enhanced co-operation and disclosure of electronic evidence (CETS No. 224), adopted in 2021.

[3] M. Anagnostopoulos and others 'Botnet command and control architectures revisited: Tor hidden services and fluxing.' (2017) in Athman Bougettaya and others (eds) *Web Information Systems Engineering–WISE 2017: 18th International Conference, Puschino, Russia, October 7-11, 2017, Proceedings, Part II* (Springer 2017), pp. 517–527.

https://doi.org/10.1016/j.clsr.2024.106080

Available online 21 November 2024

negligible compared to the prevalence of botnet activity.[4,5] Furthermore, even if a botmaster can be identified (and caught) the botnet may be reactivated by another botmaster, making it very hard to shut down a botnet irreversibly.[6]

Apart from these practical issues, botnet defence faces legal constraints on its own. This applies obviously to law enforcement, but also to protection mechanisms on the level of private actors, the latter being particularly important in botnet defence given the practical constraints of the former. Because mitigating botnets typically requires the inspection of Internet communication data in some way, a natural source of such constraints is (European) data protection law. Given the open phrasing of many of its provisions, the legal constraints set out, most notably, by the GDPR rarely provide operational rules, leading to legal uncertainty. This materializes, for instance, in the context of cybersecurity information sharing. While here one of the major concerns articulated in the industry is compliance with privacy regulations[7] it is not at all clear whether the legal framework is in fact as restrictive as perceived.[8] Somewhat paradoxically, it seems that privacy intrusiveness is not yet well established as a dimension in the design of cybersecurity techniques.[9]

These issues illustrate the necessity to contribute to the relatively sparse body of literature on data protection law and how it relates to the use of cybersecurity tools with a specific focus on botnet defence. In this particular context, there is—to the best of our knowledge—only one analysis[10] covering this specific topic in some depth. We believe that there is a need to critically assess and elaborate on this existing work. In particular, we provide a more detailed assessment of some paradigmatic botnet defence scenarios. Although it is impossible to address all the legal issues that arise, our contribution assists the development of a more nuanced discussion by connecting different contexts with different legal issues, while taking thoroughly into account the current developments in EU legislation and jurisprudence.

The remainder of this article is structured as follows. First, we sketch the high-level fundamentals of botnets and botnet defence from a technical perspective. We then introduce the relevant legal framework and discuss in more depth some issues that most botnet defence scenarios face, such as the question of whether IP addresses constitute personal data. After that, we assess four paradigmatic botnet defence scenarios. While we focus on the GDPR as the main set of rules, we also cover anecdotally adjacent legal regimes, such as the ePrivacy Directive.[11]

2. Technical background

This section gives a generic and simplified sketch of the technical background concerning the functioning of a botnet and botnet mitigation techniques.[12]

As indicated above, botnets are a collection of infected machines that are used to conduct further cyber-attacks. Cyber-attacks via botnets involve three steps: First, the attacker must gain control of a sufficient number of machines, which typically requires the installation of some malware (the distribution of which may already make use of the botnet assembled so far). Second, the attacker must be able to communicate regularly with the bots and instruct them according to the use case; this communication can either take place between the attacker's server (Command-and-Control or C2 server) and each bot directly, via intermediate platforms or servers, or the bots themselves can be used to propagate information using peer-to-peer mechanisms (thus minimizing connections to the C2 server). Finally, the bots carry out a cyber-attack according to the attacker's instructions.

Reflecting this logic, defences against botnets can address different levels of the botnet infrastructure. First, the botnet itself can be targeted by either shutting down the C2 server or disrupting its reachability for bots as far as possible. A common strategy exploits the fact that some botnets use the DNS system[13] to obtain the current IP address of the C2 server. If the domains related to the C2 server can be identified, the DNS lookup process can be modified; that is, not the C2 server's IP address is returned but the traffic is redirected to another server (sinkholing). This requires access to the DNS system and, hence, for instance, the cooperation of Internet Service Providers (ISPs). Second, botnet defence can refer to the protection on the individual level against becoming/remaining part of the botnet. This can be achieved by performing malware scans or by blocking communication with known C2 servers or bot peers. A standard technique in network security is to provide a blacklist of known C2 server IP addresses to the local firewall, blocking the communication (blacklisting); in principle, this can also be done with DNS requests, leading, in some sense, to a sinkhole at the individual level. Such lists are available open-source or from cybersecurity companies.[14] Finally, protection can be tailored to botnet-related attacks; the main problem related specifically to botnet-mediated attacks (in contrast to similar attacks performed by individual attackers directly) is its magnitude and the fact that the attack is mediated by otherwise legitimate devices, thus obfuscating the link to the malicious source.

The main challenge in applying those techniques lies at the beginning, namely in identifying the malicious sources by determining, most notably, their IP addresses and further characteristics. We focus on

[4] Notable exceptions are, for instance, the takedown of the EMOTET botnet by international cooperation coordinated by Europol and Eurojust, see Europol, 'World's most dangerous malware EMOTET disrupted through global action,' (27 January 2021) https://www.europol.europa.eu/media-press/newsroom/news/worlds-most-dangerous-malware-emotet-disrupted-through-global-action, or the takedown of Quakbot, 'Qakbot botnet infrastructure shattered after international operation', (30 August 2023) https://www.europol.europa.eu/media-press/newsroom/news/qakbot-botnet-infrastructure-shattered-after-international-operation last accessed 05 February 2024.

[5] For instance, ENISA found 17.602 functional C2 servers (2019) https://www.enisa.europa.eu/topics/cyber-threats/threats-and-trends/etl-review-folder/etl-2020-botnet last accessed 05 February 2024.

[6] L. Abrams, 'Emotet malware attacks return after three-month break' (07 March 2023) https://www.bleepingcomputer.com/news/security/emotet-malware-attacks-return-after-three-month-break/ last accessed 05 February 2024.

[7] For a comprehensive analysis of the barriers perceived, A. Zibak, A. Simpson, 'Cyber Threat Information Sharing: Perceived Benefits and Barriers' *in Ares '19: Proceedings of the 14th international conference on availability, reliability and security* (ASM 2019).

[8] C. Sullivan, E. Burger, '"In the public interest": The privacy implications of international business-to-business sharing of cyber-threat intelligence' (2017) 33 Computer Law & Security Review 14; M. Horák, V. Stupka, & M. Husák, 'GDPR compliance in cybersecurity software: A case study of DPIA in information sharing platform' *in Ares '19: Proceedings of the 14th international conference on availability, reliability and security,* (ASM 2019).

[9] For an initial framework, see Eron Toch and others, 'The privacy implications of cyber security systems: A technological survey' (2018) 51 (2) ACM Computing Surveys CSUR 1.

[10] L. Böck and others, 'Processing of botnet tracking data under the GDPR' (2022) 45 (105652) Computer Law & Security Review 1.

[11] Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) [2002] OJ L201/37.

[12] For a more detailed but still concise description of the basics, see, for example, M. Bailey and others, 'A survey of botnet technology and defenses' *in Cybersecurity Applications & Technology Conference for Homeland Security* (IEEE 2009), pp. 299–304. More recent and detailed, S. N. T. Vu and others, 'A survey on botnets: Incentives, evolution, detection and current trends', (2021) 13 (8) Future Internet 1.

[13] The DNS system translates a domain name (which is what is typed in the browse, e.g., "google.com") to an IP address associated with that URL.

[14] For instance, <https://sslbl.abuse.ch/blacklist/> last accessed 05 February 2024.

detection techniques that operate primarily on the network level. Here, we can distinguish roughly between targeted methods and monitoring methods.[15] Targeted methods try to get involved in the specific communication processes of a botnet actively. For instance, by placing vulnerable-by-design machines (honeypots) on the internet defenders hope to get infected and thus identify, analyse, and maybe even infiltrate the botnet.[16] Monitoring techniques, in turn, examine a vast amount of network traffic and perform analysis of this traffic data. For example, anomaly-based detection techniques check network traffic against a baseline and issue alerts in case of significant deviations. For this reason, passive detection techniques typically rely on a large amount of accumulated (logged) network data—such as IP addresses or DNS queries, along with additional identifiers such as timestamps—to recognise such patterns.

3. Legal framework

This section introduces the relevant legal framework for assessing particular botnet mitigation techniques. As a general rule, data protection law applies whenever personal data is processed, which in our context requires a non-trivial assessment of IP addresses (*infra 3.1.*). Processing of personal data is then constrained and accompanied by a variety of provisions, most notably the requirement of a legal basis (*infra 3.2.*) and the obligation to comply with multiple principles of processing (some of which we present *infra 3.2*).

While the GDPR is the most prominent instance of European data protection law, multiple other sets of norms can become relevant. Of importance is, firstly, the Law Enforcement Directive[17] whose application must be considered whenever law enforcement authorities process personal data for such purposes. For providers of electronic communication services—most important for us are Internet Service Providers (ISPs)—the ePrivacy Directive provides, according to Article 1(2) ePrivacy Directive, Article 95 GDPR, a (partly) more specific regime.[18] Furthermore, EU institutions are regulated by Regulation (EU) 2018/1725.[19] However, the variety of legal regimes notwithstanding, many of the main concepts overlap, in particular the notion of personal data (Article 4(1) GDPR, Article 3(1) Law Enforcement Directive, Article 3(1) Regulation (EU) 2018/1725). Without much loss of generality, hence, our analysis is for simplicity embedded in the GDPR.

Another point to keep in mind is that parts of the legislation and jurisprudence discussed here refer to the Data Protection Directive,[20] which has been the predecessor of the GDPR. In that regard, Article 94 (2) GDPR states that references to this Directive in other legal acts shall now be considered as referring to the GDPR; this holds analogously for the case law of the Court of Justice of the European Union (CJEU).[21] Thus, we implicitly transpose such references.

3.1. Involvement of personal data

The GDPR applies only to the processing of personal data, Article 2 (1) GDPR. According to the definition in Article 4(1) GDPR, personal data is "any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier […] of that natural person". As a general rule, the notion of personal data must be interpreted broadly.[22]

In general, network-based botnet defence techniques process Internet communication data. To fall within the scope of the GDPR, such data must be related to an identifiable data subject. However, for Internet traffic, the criterion of identifiability is non-trivial because the traffic is as such related only to (global) IP addresses.[23] In itself, an IP address—at least when using IPv4[24]—contains only information linking it to the managing ISP[25] and potentially some geolocation data. By managing we mean that the ISP is authorised to assign this IP address to one of its subscribers, which can be done either ad hoc (e.g., for one day; dynamic IP address) or for a longer time period (static IP address). Thus, the mapping of an IP address onto the ISP's subscriber is, in general, known only to this ISP. Furthermore, the customer is not necessarily the person using this IP address to connect to the Internet (think, for instance, of a friend using the subscriber's Wi-Fi). Finally, because of the frequent use of Network Address Translation (NAT), multiple users can share one common IP address, while the content is distributed individually at the local level; in this case, the IP address (alone) does not even uniquely relate Internet traffic to *some* user.

Therefore, it is of paramount importance for the material scope of the GDPR in our context to determine whether IP addresses are personal data[26] or suitable online identifiers for mapping other related data, such as DNS requests, onto an individual.[27] We argue that the relevant jurisprudence of the CJEU is far from straightforward as regards this question and leaves room for disqualifying IP addresses as personal data in some scenarios relevant for us.

[15] L. Böck and others, 'Processing of botnet tracking data under the GDPR' (2022) 45 (105652) Computer Law & Security Review 1, 3-4.

[16] B. Stone-Gross and others, 'Your botnet is my botnet: analysis of a botnet takeover' *in Proceedings of the 16th ACM Conference on Computer and Communications Security* (ASM 2009), pp. 635-647.

[17] 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 [2016] OJ L119/89. Recall that, in general, a directive must be transposed into the national law of the Member States in order to have legal effects (Article 288(4)) TFEU. For our analysis, we will abstract from this as far as possible and generally refer to directives as to regulations (which are directly applicable).

[18] See, to that effect, Joined Cases C-511/18, C-512/18 and C-520/18, *La Quadrature du Net v. Permier ministre*, ECLI:EU:C:2020:791, para 202. For more details on the interplay between the two regimes, see EDPB, 'Opinion 5/2019 on the interplay between the ePrivacy Directive and the GDPR, in particular regarding the competence, tasks and powers of data protection authorities' (12 March 2019) < https://www.edpb.europa.eu/sites/default/files/files/file1/201905_edpb_opinion_eprivacydir_gdpr_interplay_en_0.pdf> last accessed 25 October 2024.

[19] Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC [2018] OJ L295/39.

[20] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data [1995] OJ L281/31.

[21] Case C-597/19 *M.I.C.M. v. Telenet BVBA* ECLI:EU:C:2021:492, para 107.

[22] See, e.g., Case C-180/21 *Inspektor v Inspektorata kam Visshia sadeben savet* ECLI:EU:C:2022:967, para 70.

[23] (Global) IP addresses are unique identifiers across the Internet as opposed to local IP addresses, which are assigned by the administrator of a local network and are valid only within this local network.

[24] For IPv6, it depends on the implementation. Potentially, although not favorably, IPv6 addresses potentially disclose (parts of) the MAC address of the physical device used to access the Internet.

[25] We do not consider that other entities could potentially obtain IP addresses from Internet registries and thus be related to the IP address directly. For private persons, this is a rather hypothetical scenario anyway.

[26] See L. Böck and others, 'Processing of botnet tracking data under the GDPR' (2022) 45 (105652) Computer Law & Security Review 1, 4

[27] Although analytically distinct, both questions depend on whether a sufficient connection to an identifiable natural person can be established, whence we treat them uniformly.

3.1.1. Relevant jurisprudence of the CJEU

With regard to the quality of (global and dynamically assigned) IP addresses as personal data, the CJEU's leading case law states that in case the processing entity itself cannot map an IP address onto a particular individual the IP address must be considered personal data in relation to the entity concerned if this entity "has the legal means which enable it to identify the data subject".[28] The case concerned, in short, a server retaining IP addresses along with timestamps for cybersecurity and, potentially, law enforcement purposes. A client, who had connected to the server using a dynamically assigned IP address *and* who was also the subscriber to the internet access point complained about the retention on the grounds of data protection law. For the assessment of whether the IP address retained was personal data, the CJEU assumed that no further identifying information was available to the server retaining the IP address,[29] whence that server was not by itself able to map the IP address onto the user. But according to the CJEU, the identifiability criterion is fulfilled (also) with respect to the server if there are legal means available to that server that provide access to individualizing data in the hands of another entity. Since the CJEU furthermore assumed(!)[30] that suitable information was in the hands of the ISP (which according to the facts of the case was true because of the coincidence between ISP customer and user) the Court concluded that if the national law provides for means to acquire that information from the ISP the IP address must be considered personal data for the entity operating the server. In effect, the CJEU thus relativised a prior decision, *Scarlet Extended*, according to which IP addresses—without any discussion and in the context of the Charter[31]—must be considered personal data (categorically).[32] The Court justified a distinct assessment with, first, the *perspective* of the controller (in *Scarlet Extended*, the controller was an ISP) and, second, the modality of the IP address assignment (which was static in *Scarlet Extended*).[33]

Despite the narrow scope of the case,[34] the Breyer formula has been since used as a standard in the CJEU case law to assess identifiability whenever third-party support is required for identification. Although not directly related to IP addresses, the subsequent decisions can be regarded as confirmations as well as particularizations of the Breyer formula and thus backpropagate to the assessment of IP addresses.

The formula was confirmed in *M.I.C.M*,[35] which concerned the registration of IP addresses for the purpose of subsequent use in legal proceedings by private actors. The formula was then extended to comprise other intrinsically meaningless identifiers, namely vehicle chassis identifiers (VIN), which individualize vehicles (and, potentially, their owners). According to the CJEU, such VINs are not necessarily personal data, depending on the additional information acquirable to the entity in question.[36] The Advocate General in the underlying opinion explicitly distinguished this case from another VIN-related procedure in which public institutions requested access to a VIN database; since the public authority in the other case was able to access the public vehicle registration files it could also infer the identity of the persons who registered the vehicle. For other entities, the accessibility of such information must be determined according to the circumstances of the individual case.[37]

Furthermore, the Breyer formula has been invoked in the assessment of whether publishing a press release constitutes processing of personal data when the individual referred to by the press release can, from the perspective of the addressees of the press release, only be inferred by external means.[38] In this case (*OC/Commission*), the CJEU, referred to factual (as opposed to legal) means *from the perspective of the recipient* (i.e., not the controller)[39] of the information. The Court concluded that the fact that an investigative journalist eventually identified the natural person to which the press release referred was as such insufficient to justify identifiability,[40] indicating that not all possible configurations of how a person can be defined need to be taken into account. While the Court did not explicitly confirm the General Court's criterion of an "average reader" for objectively determining the reasonableness (and, thus, significance) of reasonable means to be used for identification,[41] it did not declare its restrictive reasoning as invalid per se; it only opposed the General Court's conclusion that the presence of specialized journalists in a case concerning fraud is not negligible, whence the risk of identification, given external sources of information on the web, was not insignificant (which is quite obvious).[42]

In a case currently under appeal, the General Court extensively relied on *Breyer* to assess whether the disclosure of pseudonymized documents constitutes processing of personal data and concluded that this is not the case when the receiving parties are reasonably unable to access the pseudonymization mapping.[43] It remains to been seen whether this reasoning will be upheld by the CJEU.

It should be finally noted that the CJEU—in the spirit of *Scarlet Extended*—very recently stated that IP addresses are personal data according to its case law without mentioning any restrictions.[44] However, this statement was rather anecdotical and the CJEU pointed explicitly to the *M.I.C.M.* case for further reference. Therefore, albeit irritating, it cannot be assumed that the Court intended to change the assessment towards declaring IP addresses to be personal data categorically (again).

3.1.2. Discussion

There are a couple of things worth highlighting, some of which remain controversial within the academic discussion of the outlined case law. We first discuss the implications on a general level before drawing implications for our context.

First and most importantly, the CJEU assesses the identifiability of a natural person generally from the perspective of the entity concerned, i.e., only considers identifying information (potentially) available to that

[28] Case C-582/14 *Breyer v Bundesrepublik Deutschland* ECLI:EU:2016:779, para 49.

[29] Case C-582/14 *Breyer v Bundesrepublik Deutschland* ECLI:EU:2016:779, para 37.

[30] Case C-582/14 *Breyer v Bundesrepublik Deutschland* ECLI:EU:2016:779, para 37: "The referring court's first question is *based therefore on the premiss* […] that […] the internet services provider has additional data which, if combined with the IP address would enable the user to be identified."

[31] Case C-70/10 *Scarlet Extended v. SABAM* ECLI:EU:C:2011:771, para 51.

[32] Charter of Fundamental Rights of the European Union [2012] OJ C326/391.

[33] Case C-582/14 *Breyer v. Bundesrepublik Deutschland* ECLI:EU:2016:779, paras 34-36.

[34] The AG extensively elaborated on the peculiarities of the question referred, Case C-582/14 *Breyer v Bundesrepublik Deutschland* ECLI:EU:C:2016:339, Opinion of AG Sánchez-Bordona, paras 45 et seq.

[35] Case C-597/19 *M.I.C.M v Telenet BVBA* ECLI:EU:2021:492, para 102.

[36] Case C-319/22 *Gesamtverband Autoteile-Handel eV v Scania* ECLI:EU:2023:837, para 49.

[37] Case C-319/22 *Gesamtverband Autoteile-Handel eV v Scania* ECLI:EU:C:2023:385, Opinion of AG Sánchez-Bardona, paras 29 et seq.

[38] Case C-479/22 P, *OC v Commission* ECLI:EU:C:2024:215.

[39] This is not taken into account by A. Lodie, "Case C-479/22 P, Case C-604/22 and the limitation of the relative approach of the definition of 'personal data' by the ECJ", <https://europeanlawblog.eu/2024/03/25/case-c-479-22-p-case-c-604-22-and-the-limitation-of-the-relative-approach-of-the-definition-of-personal-data-by-the-ecj/>, who claims that the Court adopted a absolute approach.

[40] Case C-479/22 P *OC v Commission* ECLI:EU:C:2024:215, para 58.

[41] Case T-384/20 *OC v Commission* ECLI:EU:T:2022:273, para 76: "lecteur moyen".

[42] Case C-479/22 P *OC v Commission* ECLI:EU:C:2024:215, para 57.

[43] Case T-557/20 *SRB v. EDPS* ECLI:EU:T:2023:219.

[44] Case C-470/21 *La Quadrature du Net v Permier ministre* ECLI:EU:C:2024:370, para 60.

entity ("in relation to that provider").[45] Furthermore, for the processing operation of disclosure, the perspective of the recipient(s) is decisive. This means that the CJEU advocates a so-called relative approach to identifiability as opposed to an absolute one, where according to the latter it is sufficient for a natural person to be identifiable that *some* entity can identify the individual.[46] Related to this is the implication that, arguably, the mere singling out of an individual in the sense that, disregarding NAT, there is a unique mapping from IP address to *some* user does not establish identifiability.[47] This is because, again disregarding NAT, the dynamic IP address (along with a timestamp) objectively determines a particular user *at some point in time* (no less than a static IP address). Apparently, there must be some deeper identification of the individual to meet the threshold of identifiability, the concrete shape of which remains unclear. This is also confirmed, e.g., by the *OC/Commission* case, where the press release as such objectively referred to a particular individual, whereas the CJEU still assessed the means to (further) identify that individual.

The academic literature, to be sure, has argued that the Court in *Breyer* either had not touched upon the singling-out problem[48] or that there is nevertheless room for a singling-out approach.[49] The latter approach stresses the fact that *Breyer* was about the *retention* of an IP address, which does not preclude that *during the ongoing session* the user is "reached" and thus identified by the IP address as "a flesh and blood individual"[50] due to the ongoing interaction. Such constructions highlight that a straightforward understanding of *Breyer* can lead to defective protection of subjects in other contexts, such as online tracking for advertisement purposes.[51] When taken as a general rule for the assessment of IP addresses, however, they circumvent the CJEU's understanding of the limitations of the scope of the GDPR, at least when reaching out shall be understood as the process of exchanging data, which is the whole purpose of IP addresses. Besides, when delimiting personal data temporarily by an ongoing social interaction, it is unnecessary to rely on IP addresses as personal data since the data of interest, namely the data from the interaction itself along with, e.g. a session cookie, can be considered personal; this would furthermore overcome the problem related to the fact that the IP address is only indirectly relating to the user (namely, via the ISP, its customer and, possibly, the NAT system).

Secondly, according to the CJEU, it seems to suffice for available legal means in the sense of the Breyer formula that they are conditioned on a chain of future and uncertain events. Even though the legal means in the *Breyer* case stemmed from German law and thus were not for the Court to interpret, the Court acknowledged that "it seems however, subject to verifications to be made in that regard by the referring court that, in particular, in the event of cyber attacks legal channel exist"[52] and therefore "it appears that the online media service provider has the means which may likely reasonably be used in order to identify the data subject".[53] It should be noted that those legal channels in German law require a criminal proceeding, within which the legal enforcement authorities can obtain information from the ISP. For a non-malicious user, thus, the probability of being identified is negligible in practice. While the referring court in the judgment following the CJEU`s decision without further argumentation assumed that the respective means are reasonably likely to be used even for a non-malicious party,[54] this hardly matches the cautious wording of the CJEU. When read in connection with the Court's statement according to which the identification is not reasonably likely "if the identification of the data subject was prohibited by law or practically impossible on account of the fact that it requires a disproportionate effort in terms of time, cost and man-power, so that the risk of identification appears in reality to be insignificant" ,[55] practical feasibility remains a substantial constraint to be taken into consideration when assessing a particular case.[56] Put differently, even though the threshold of reasonable likelihood might not be very strict,[57] it should nevertheless be marginalized completely. Such an interpretation is supported by the subsequent decisions, namely *OC/Commission*, where the Court extensively investigated whether the risk of the circumstances which led to the identification of the person referred to by the press release was ex ante insignificant. What then reasonably likely means in detail, is yet to be determined by future case law. In any case, a purely hypothetical legal channel should not be regarded as sufficient.

Thirdly, it remains an open question how to deal with the CJEU`s distinction regarding static and dynamic addresses as well as the perspective of an ISP versus a non-ISP (meaning entities that are not ISPs). From the perspective of the Breyer formula, there is, in principle, no reason to treat static and dynamic IP addresses differently, for both variants require additional information to identify the user; it is only that the risk of identification based on accumulated information (e.g., based on web tracking) is higher for static addresses.[58] As regards ISPs, it should be recalled first that the CJEU`s reasoning in *Breyer* is based on the assumption that the (associated) ISP can identify the *user*,[59] which is not true in general[60] and it seems as a matter of fact even questionable

[45] C-582/14 *Breyer v Bundesrepublik Deutschland* ECLI:EU:2016:779, para 49, where the provider is the server, i.e., the controller.

[46] For a discussion of this issue, M. Finck and F. Pallas, 'They who must not be identified—distinguishing personal from non-personal data under the GDPR' (2020) 10 International Data Privacy Law, 11, 17-18.

[47] See also, including a broader discussion of the academic literature on the singling out approach, P. A. E. Davis, 'Facial Detection and Smart Billboards: Analysing the "Identified" Criterion of Personal Data in the GDPR' (2020) 6 European Data Protection Law Review, 365, 372 et seq.

[48] F. Z. Borgesius, 'The Breyer Case of the Court of Justice of the European Union: IP Addresses and the Personal Data Definition' (2017) 3 European Data Protection Law Review, 130.

[49] N. Purtova, 'From knowing by name to targeting: the meaning of identification under the GDPR' (2022) 12 International Data Privacy Law, 163.

[50] N. Purtova, 'From knowing by name to targeting: the meaning of identification under the GDPR' (2022) 12 International Data Privacy Law, 163, 179.

[51] F. J. Z. Borgesius, 'Singling out people without knowing their names – Behavioural targeting, pseudonymous data, and the new Data Protection Regulation' (2016) 32 Computer Law and Security Review, 256. It seems that even the CJEU in a case concerning web advertisement assumes that if one knows the IP address of the "user" means that one can identify/ profile that user, see Case C-604/22, *IAB Europe v Gegevensbeschermingsautoriteit*, ECLI:EU:C:2024:214, para 44.

[52] Case C-582/14 *Breyer v Bundesrepublik Deutschland* ECLI:EU:2016:779, para 47.

[53] Case C-582/14 *Breyer v Bundesrepublik Deutschland* ECLI:EU:2016:779, para 48.

[54] Bundesgerichtshof, VI ZR 135/13, ECLI:BGH:2017:160517UVIZR135.13.0, para 26. For a (short) comparision between the referring court and the CJEU in that regard, reaching a similar conclusion, U. Baumgartner, 'Sind IP-Adressen wirklich immer personenbezogene Daten? Ein Zwischenruf' (2023) 13, Zeitschrift für Datenschutz, 125, 126.

[55] Case C-582/14 *Breyer v Bundesrepulik Deutschland* ECLI:EU:2016:779, para 46.

[56] In this sense, F. Z. Borgesius, 'The Breyer Case of the Court of Justice of the European Union: IP Addresses and the Personal Data Definition' (2017) 3 European Data Protection Law Review, 130, 136.

[57] M. Finck and F. Pallas, 'They who must not be identified–distinguishing personal from non-personal data under the GDPR' (2020) 10 International Data Privacy Law, 11, 18.

[58] Article 29 Working Party, 'Opinion 04/2007 on the Concept of Personal Data (WP 136) 01248/07/EN', p. 16.

[59] E.g. Case C-582/14, *Breyer v Bundesrepublik Deutschland* ECLI:EU:2016:779, para 39: "user of a website".

[60] See also, P. A. E. Davis, 'Facial Detection and Smart Billboards: Analysing the "Identified" Criterion of Personal Data in the GDPR' (2020) 6 European Data Protection Law Review, 365, 373.

whether this can be said to be true heuristically.[61] What can be said, though, is that the ISP can identify its *subscribers*. In that regard, now under the obvious constraint that the subscriber is a natural person, an IP address is a piece of personal data from the perspective of the ISP to which the IP address is associated. Thus, while the content of the information changes from relating Internet traffic to the individual causing this traffic to relating Internet traffic to the subscriber of an access point who only possibly causes the observed traffic, the GDPR is applicable. With this perspective, *Breyer* can be both based on solid grounds, that is, without making a non-realistic assumption, and reconciled with *Scarlet Extended*, in which case it was claimed that IP addresses generally are personal data, to the extent that the latter judgment concerned an ISP (in charge of assigning the respective IP addresses).

3.1.2. Specific consequences for botnet defense techniques

As for the implications for our context, the first and most important observation from the point of view of the settled CJEU case law is that the assessment of botnet defense techniques requires a case-specific analysis of the entity and its informational background. In particular, the claim that processing IP addresses implies processing of personal data[62] is not justified in general. This is particularly true in cyber defense contexts, as they comprise a variety of actors deploying very different techniques with different objectives.

More concretely, the case law requires, as a first step, a distinction to be made between an ISP and a non-ISP, since IP addresses are, at least when one sticks closely to the jurisprudence of the CJEU, personal data for an ISP categorically. It should be noted that since ISPs are not in a privileged position with respect to the IP addresses they do not assign, this conclusion is not justified regarding IP addresses governed by other ISPs; insofar, we must treat the ISP as a non-ISP.

Insofar as an entity deploys a defense technique, most notably a monitoring technique, that inspects *local* IP addresses, the informational background of this entity compares to that of an ISP in that both control the assignment of the IP addresses.[63] In fact, the assignment on the local level is oftentimes more fine-grained, for this assignment typically relates the IP addresses more directly to individuals (for instance, via an authentication mechanism). The *Breyer* formula straightforwardly comprises such cases as processing of personal data. From an analytical perspective, therefore, the distinction between ISPs and non-ISPs turns out to be less important for the purposes of determining whether a piece of data is personal than whether this data constitutes *internal* or *external* data for this entity.

For external IP addresses now, it must be asked whether an entity can link the IP address processed to a natural person. If the entity cannot map the IP address themselves, it depends on whether they can identify the data subject through the ISP or the information of another entity by means reasonably likely to be used. In the context of botnet defense, there are multiple issues that limit the possibility to obtain information on the user of an IP address. As regards connections with C2 servers, it should be noted that those servers are typically located in countries outside the EU,[64] impeding the access to identifying information or even rendering it (practically) impossible. With respect to connections with peer bots (in a distributed setting), non-ISPs may not have means to identify the natural person behind the peer bot since this natural person generally is not malicious, which will typically lead to a lack of legal channels via law enforcement.

Additionally, observe that in many cases of botnet defense conducted by private entities, the primary objective is not to initiate public prosecution or even enforce potential civil claims, but simply to protect the own network against attacks. A prototypical example is a traditional firewall, which simply blocks a connection. This leads to the question of whether the purpose of the processing is a relevant criterion in assessing identifiability. For the converse case, the Art. 29 Working Party has argued that when the purpose of processing is identification, it cannot be coherently said that the data subject is not identifiable.[65] When considering this from the point of view of the CJEU's risk criterion, it is consequent to reverse this argument: when the data of a particular technique shall not be used for identification the risk of identification is at least lowered. In connection with further aspects, such as difficulties to map the IP address, this can lead to an overall negligible risk of identification. To compare, note that the data controller in *Breyer* collected the IP addresses explicitly in order to enable public prosecution.[66]

Overall, we observe that already the *Breyer* case is far from postulating that IP addresses are always personal data. In the current state of the jurisprudence the relevant criterion is whether the controller has either legal or (other) reasonable means at their disposal to resolve the IP address. A major aspect for determining such means is whether the data concerned should be considered internal or external with respect to the potential controller. For internal data, the identifiability criterion is trivial with respect to the (natural) person linked to the IP address in the entity's records. Regarding external data, a careful examination is required taking into account, in particular, the specific defense technique and the circumstances in which it is deployed.

3.2. Legal basis

Every processing of personal data requires a legal basis (Article 5(1)(a) GDPR). Article 6(1) GDPR provides six options which we can very roughly categorise into consent-based, obligation-based, and interest-based, respectively.[67] In the following, we give an overview and discuss aspects relevant for botnet mitigation in general, while we elaborate on specific aspects and problems later. Note that since the processing of special categories of personal data pursuant to Article 9(1) GDPR adds another layer of complexity, we focus solely on non-sensitive personal data.

3.2.1. Consent-based

A consent-based legal basis relies on the autonomous opting-in of the data subject. The consent can concern the processing of data specifically (Article 6(1)(a) GDPR) or be implicit by contracting with the controller (Article 6(1)(b) GDPR). It is important to see that this legal basis is suitable, if at all, only for the data of subjects the controller has an individual relationship with. Therefore, if personal data of a third person is processed we (additionally) require a different legal basis.

3.2.2. Obligation-based

According to Article 6(1)(c) GDPR, the processing of personal data is legitimate if it is necessary to fulfil a legal obligation of the controller. In

[61] Considering this is P. A. E. Davis, 'Facial Detection and Smart Billboards: Analysing the "Identified" Criterion of Personal Data in the GDPR' (2020) 6 European Data Protection Law Review, 365, 374.

[62] At least implicitly, L. Böck and others, 'Processing of botnet tracking data under the GDPR' (2022) 45 (105652) Computer Law & Security Review 1, 5.

[63] Similarly, F. Z. Borgesius 'The Breyer Case of the Court of Justice of the European Union: IP Addresses and the Personal Data Definition' (2017) 3 European Data Protection Law Review, 130, 136.

[64] ENISA, 'Botnet. ENISA Threat Landscape. From January 2019 to April 2020', (2020) <https://www.enisa.europa.eu/topics/cyber-threats/threats-and-trends/etl-review-folder/etl-2020-botnet>, p. 12, last accessed 05 February 2024.

[65] Article 29 Working Party, 'Opinion 04/2007 on the Concept of Personal Data (WP 136) 01248/07/EN', p. 16.

[66] Case C-582/14 *Breyer v. Bundesrepublik Deutschland* ECLI:EU:2016:779, para 14.

[67] Note that this grouping is not established and serves only to structure our analysis.

our context—and restricting ourselves to potential legal obligations on the EU level—the primary candidates for such legal obligations are provisions that require entities to ensure, broadly speaking, cybersecurity. Consider, for instance, Article 32(1)(b) GDPR,[68] which essentially requires controllers to ensure that the processing of personal data takes place within secure systems. Other provisions that require cybersecurity (and are not linked to data protection) are contained in sector-specific regulation, in particular with regard to essential and important entities (Article 21(1) NIS 2 Directive[69]) or for ISPs (Article 4(1) ePrivacy Directive); these provisions, as a general rule, must be transposed into national law first (Article 288(3) TFEU) to establish a legal obligation.

As regards potential conflicts with data protection, it is important to note that at least the provisions in the ePrivacy and NIS 2 Directives must be understood without prejudice to data protection rules (Article 2 (14) and Recital 14 NIS 2 Directive, Article 4(1a) ePrivacy Directive), meaning that they cannot justify any processing of personal data without checking its conformity with the GDPR. In this sense, Recital 121 of the NIS 2 Directive states that Article 6(1)(c) GDPR "could be considered" a suitable legal basis for the "processing of personal data, to the extent necessary and proportionate". More generally, since personal data is not involved in all possible cybersecurity measures, data protection rules are more specific. For the transposition into national law, it is thus up to national legislation to incorporate, in respect of Article 6(3) GDPR, aspects of data protection in the obligations such that those can constitute an obligation in the sense of Article 6(1)(c) GDPR.

3.2.3. *Interest-based*

As regards interest-based clauses, we have for public interest tasks and official authorities Article 6(1)(e); for private controllers Article 6 (1)(f) GDPR provides the main legal basis. Article 6(1)(e) GDPR requires that the processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority. Article 6(1) (f) GDPR refers to a *legitimate interest* of the controller and the *necessity* and *proportionality* of the processing.[70] Given the particular importance that European legislation attaches to it (Recital 3 NIS 2 Directive), cybersecurity can be considered both a public interest and a legitimate interest of the controller. For public interest, Recital 3 NIS 2 Directive refers to "damages to the Union's economy and society" and thus the essential role of cybersecurity for the prosperity of the internal market; more concretely, Recital 121 NIS 2 Directive stipulates that the public interest clause could be used to justify the processing of personal data by certain entities. Regarding the legitimate interest (of private entities), the GDPR itself refers to this clause when considering processing by certain entities, such as CSIRTs or cybersecurity service providers (Recital 49 GDPR).[71] Furthermore, the CJEU explicitly considers network security as a legitimate interest of a controller even beyond entities mentioned in the recital.[72]

Concerning necessity, it can be generally observed that technical measures are essential in combating botnets.[73] As to which technique to implement concretely, the necessity criterion constrains the controller to select the least intrusive one among those similar in effectiveness, in particular as regards the requirement to minimise the data processed.[74]

The main restriction comes from the proportionality requirement. Even though data protection and cybersecurity are to a large degree complementary—as witnessed, for instance, by Article 32 GDPR—they also can conflict.[75] In such cases, we must weigh the legitimate interest of the controller and the impact on the data subject and compare both. This requires a case-by-case assessment, but we can make some general observations.[76]

On the one hand, the interest in ensuring cybersecurity is generally of a high magnitude. This follows not only from the perspective of the EU legislator as documented by the aforementioned Recital 3 NIS 2 Directive but also from the fact that this Directive has been substantially broadened both in terms of scope and substance compared to the predecessor directive,[77] indicating an increase in the relevance of cybersecurity at the regulatory level. This is confirmed by the recent adoption of further cybersecurity-related legislation, such as the Cybersecurity Act.[78] Given that a botnet infrastructure is considered the backbone of modern cybercrime,[79] the relevance of cybersecurity transposes directly to botnet defence. Furthermore, it should be noted that botnet defense, at least on the network level, typically requires the analysis of some (potentially) personal data, most notably IP addresses. Now, it has been argued—in the vein of the CJEU[80]—that the processing of an IP address can be justified in particular when the interest pursued cannot be otherwise satisfied at all.[81] Assuming that botnets can be detected mainly on the network level, this also highlight the processing interest.

On the other hand, the impact on the data subject varies substantially across different techniques. The main criterion in this respect is, again from the perspective of data minimisation, the density of the information captured, that is, how much information about the data subject can

[68] In that regard, the question arises whether the GDPR can even require the processing of personal data, which A. Cormack, 'Processing data to protect data: Resolving the breach detection paradox' (2020) 17 (2) SCRIPTed, 197, calls a "paradox".

[69] Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972 and repealing Directive (EU) 2016/1148 [2022] OJ L 333/80.

[70] See, e.g. Case C-252/21 *Meta v. Bundeskartellamt* ECLI:EU:C:2023:537, paras 105-126.

[71] Note that M. D. Cole and S. Schmitz, 'The interplay between the NIS Directive and the GDPR in a cybersecurity threat landscape' (2020) University of Luxembourg Law (Working Paper No. 2019-017) <https://ssrn.com/abstract=3512093> or <https://doi.org/10.2139/ssrn.3512093>, *infra III.3,* seem to argue that the fact that other actors are not mentioned in the recital means that they cannot rely on the interest clause. We do not share this view given the generally anecdotal nature of the recital.

[72] Case C-252/21 *Meta v Bundeskartellamt* ECLI:EU:C:2023:537, para 119.

[73] J. K. Haner and R. K. Knake, 'Breaking botnets: A quantitative analysis of individual, technical, isolationist, and multilateral approaches to cybersecurity' (2021) 7 (1) Journal of Cybersecurity 1.

[74] European Data Protection Board, 'Guidelines 1/2024 on processing personal data based on Art. 6(1)(f) GDPR', <https://www.edpb.europa.eu/system/files/2024-10/edpb_guidelines_202401_legitimateinterest_en.pdf> (2024), para 29.

[75] A. Cormack, 'Processing data to protect data: Resolving the breach detection paradox' (2020) 17 (2) SCRIPTed 197, 198.

[76] For a general methodology, see also European Data Protection Board, 'Guidelines 1/2024 on processing personal data based on Art. 6(1)(f) GDPR', <https://www.edpb.europa.eu/system/files/2024-10/edpb_guidelines_202401_legitimateinterest_en.pdf> (2024), paras 31 et seq.

[77] A simple word count on the pdf files unearthed an increase from 15,237 to 38,859 words.

[78] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act) [2019] OJ L151/15.

[79] K. e Silva, 'How industry can help us fight against botnets: notes on regulating private-sector intervention' (2017) 31 International Review of Law, Computers & Technology, 105, 106.

[80] Regarding the need to process IP addresses to detect child pornography, Joined Cases C-511/18, C-512/18, C-520/12 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2020:791, para 154.

[81] Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2022:838, Opinion of AG Szpunar, paras 78 et seq.; C-470/21, *La Quadrature du Net v Premier ministre* ECLI:EU:C:2023:711, Opinion of AG Szpunar, para 58 et seq.

be inferred.[82] In its jurisprudence concerning the obligation of ISPs to retain communication data for the purposes of (potential) criminal investigations, the CJEU is particularly concerned by the risk of establishing a profile of data subjects and, hence, the severity of the interference with fundamental rights.[83] Such risk leads, according to the CJEU, to a chilling effect towards the use of telecommunication[84] and incurs the danger of data misuse.[85] Interestingly for us, the Court substantiates the risk of establishing a profile by referring, among other things, to the possibility of tracking the clickstream based on IP addresses.[86] However, the CJEU has also recently ruled that the retention of IP addresses does not per se constitute a severe interference with individual rights; for this it must be ruled out that precise conclusions about the person concerned can be made, most notably by means of a "watertight" technical and organisational separation.[87] While not directly applicable to our topic, mainly due to the law enforcement context in which this jurisprudence is embedded, and the consequences for other areas of the law still being unclear,[88] we can infer from this jurisprudence some important benchmarks for the assessment of proportionality in general.

Given those guidelines, it can be observed that a major distinction should be again made between internal and external data. For the first thing, it is much less likely that the natural person behind an external IP address is identified, even if the risk of identification is not negligible and therefore the criterion of identifiability within the notion of personal data is met.[89] In contrast, the internal perspective typically allows to map traffic onto an individual directly (at least the individual subscribing to the Internet access point, as in the case of an ISP) and hence use this information against this individual. Therefore, the risk of misuse is substantially higher for internal data. Related to this is that the internal perspective is more comprehensive in the sense that the global behaviour of the IP address on the network is visible, which can allow to establish a personal profile.[90] In contrast, observing an external IP address will only allow inferences to be made about that IP address's traffic with the controller's network, which will typically only be a subset of the overall behaviour.

Another obvious factor as regards the potential for misuse and the risk of profiling is what kind of data exactly and how much of it is inspected. A useful difference lies in whether traffic data (i.e., "data processed for the purpose of the conveyance of a communication", Article 2(b) ePrivacy Directive) or content data (i.e., the content of the communication) is inspected. While already performing traffic analysis—i.e., trying to infer information based on visible patterns in the (encrypted) network data—might reveal relatively precise information about the content of the respective traffic,[91] scanning the content of a message is typically even more invasive since it may directly reveal sensitive information. But also among traffic data it must de differentiated. For example, evaluating connections based solely on IP addresses may differ from tracking DNS queries in that a DNS query contains, in general, more granular information as it specifies the particular domain to be contacted (as opposed to just the IP address associated with the server hosting the domain). Equivalently does it make a substantial difference whether a monitoring technique tracks the partner or the type of a connection (e.g., email or web) or simply observes the unspecified activity of an IP address only.[92]

A further criterion is the amount of data subjects affected. The CJEU has also taken this into account in the context of indiscriminate data retention, with an unconstrained retention facing a substantially heavier burden of justification than when the retention is linked to the purpose of processing by some relevant and limiting criterion.[93] In this regard, note that targeted intelligence, such as setting up honeypots, does not require the indiscriminate collection of traffic data on a large scale, unlike untargeted monitoring techniques. Thus, targeted techniques are less intrusive since they minimise the processing of data and, in particular, of traffic unrelated to a botnet.[94]Finally, Recital 47 GDPR determines to take into account in balancing interests whether the data subject can reasonably expect the processing of her data, which applies rather to well-established techniques than to current research.

3.3. General principles

The GDPR defines in Article 5 the principles to be respected when processing personal data. Many of those principles are further specified by other GDPR provisions. Here, we only give a brief overview of some of the principles and elaborate on certain aspects later.

At the outset, note that insofar we are considering the realisation of those principles by technical (or organisational) means, the controller must choose measures that are appropriate, taking into account, among other things, the cost of implementation and the risk of processing (Article 25(1) GDPR). Effectively, the controller must perform a proportionality test, in which a fair balance must be stroke between the requirements of data protection and, most notably, the costs of implementation.[95] According to the CJEU, the controller benefits from some discretion which, however, is subject to a detailed judicial review.[96] Therefore, the controller should perform a structured analysis[97] and rely on a fine-grained documentation.

The principle of *data minimisation* (Article 5(1)(c) GDPR) requires that no more data is processed than necessary for the particular use case. This also includes that the data is processed in a form as protected as possible, which includes, among other things, pseudonymization

[82] To that effect, Case C-746/18 *H.K. v Prokuratuur* ECLI:EU:C:2021:152, para 34.

[83] Joined Cases C-511/18, C-512/18 and C-520/18 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2020:791, para 117.

[84] Joined Cases C-511/18, C-512/18 and C-520/18*La Quadrature du Net v Premier ministre* ECLI:EU:C:2020:791, para 118.

[85] Joined Cases C-511/18, C-512/18 and C-520/18 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2020:791, para 119.

[86] Joined Cases C-511/18, C-512/18 and C-520/18*La Quadrature du Net v Premier ministre* ECLI:EU:C:2020:791, para 153.

[87] Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2024:370, para 83 et seq.

[88] See, to that effect, Case C-597/19, *M.I.C.M. v Telenet BVBA* ECLI:EU:C:2020:1063, Opinion of AG Szpunar, para 103, pointed to a tension between the restrictive tendencies in *La Quadrature du Net* and IP address retention for the purposes of ecnforcing IP rights (para 84) and observed, more generally, that the standards set out in *La Quadrature du Net* will be hard to ignore "in other fields, such as the protection under civil law of the rights of others".

[89] Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2024:370, para 72, the CJEU stresses the proximity between IP address and identification data.

[90] Notice that the protection against profiling is of particular relevance in the GDPR. This is related to the risk of establishing a profile as understood by CJEU (e.g., Cases C-511/18 and others, *La Quadrature du Net v Premier ministre* ECLI:EU:C:2020:791, para 117).

[91] See, e.g., R. Schuster, V. Shmatikov, and E. Tromer, 'Beauty and the burst: Remote identification of encrypted video streams' in *USENIX Security Symposium* (2017) pp. 1357–1374, who shows that it is possible to infer a streamed film from packet sizes.

[92] In that sense, Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2024:370, para 80.

[93] Joined Cases C-511/18, C-512/18 and C-520/18 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2020:791, point 119.

[94] C. Sorge, 'IT security measures and their relation to data protection' G. Borges and C. Sorge (eds.), *Law and Technology in a Global Digital Society* (Springer 2021), 179, 190.

[95] Case C-340/21 *VB v Natsionalna agentsia za prihodite* ECLI:EU:C:2023:353, Opinion AG Pitruzella, para 36 (concerning primarily Art. 32 GDPR).

[96] Case C-340/21 *VB v Natsionalna agentsia za prihodite* ECLI:EU:C:2023:986, paras 43, 45, 46, 57 (concerning primarily Art. 32 GDPR).

[97] In that sense, Case C-340/21 *VB v Natsionalna agentsia za prihodite* ECLI:EU:C:2023:986, para 43: "complex assessment carried out by the controller".

(Article 25(1) GDPR). As a measure to limit the intrusiveness of the processing and thus as an aspect of data minimisation qualifies also a distributed storage of the data (in internal contexts), where, for instance, the monitored traffic data is kept apart from the data linking the monitored machine to an individual; the CJEU has explicitly required the preventive data retention for law enforcement purposes of IP addresses by ISPs to be backed by "watertight" technical solutions implementing such separation to mitigate the severity of the processing.[98]

Storage limitation according to Article 5(1)(e) GDPR captures the idea of data minimisation in a time dimension[99] and requires that personal data is deleted when no longer needed for the purpose pursued. As an example, observe that training data for ML-based anomaly detection tools can arguably be erased after training.[100] This principle is particularly difficult to handle whenever the accumulation of data strictly—albeit maybe only marginally—increases the level of security. Think of a very long retention period for forensic purposes.

As far as *purpose limitation* (Article 5(1)(b) GDPR) is concerned, the primary requirement for controllers is to define the purpose of processing ex-ante in a "specific" and "explicit" form (whenever the law does not define this purpose).[101] Since fixing a purpose frames the further assessment of the processing, in particular by providing a reference point for assessing necessity, the principles of data minimisation and storage limitation argue in favour of a precise description of this purpose, as opposed to generic descriptions such as ensuring cybersecurity.[102] As we have seen, already the intention to resolve an IP address at some point can make a difference in whether this IP address must be qualified as personal data, which also argues in favour of a careful assessment. Without prejudice to Article 6(4) GDPR, data protection does not justify the processing of data for other purposes than for the purposes for which is was collected.[103]

Furthermore, data must be processed in a *transparent* manner (Article 5(1)(a) GDPR), which requires the controller to inform the data subject about the processing and to provide further information in accordance with Article 12 et seq. GDPR. Article 14(5)(b) GDPR contains an exception for cases in which the provision of such information proves impossible or would involve a disproportionate effort; this exception becomes relevant for the processing of external traffic data. For internal data, in turn, informing the affected individuals might not only serve to comply with the transparency rules, but also, provided that the information is specific enough, to make the data processing more foreseeable and hence, as we have seen, potentially less intrusive.

Finally, the *integrity and confidentiality* of the data must be ensured according to Article 5(1)(f) GDPR by means of appropriate technical and organizational measures. This can be understood as a derived principle to safeguard other principles by means of an adequate framework for the processing.[104] Consider, for instance, antivirus software used to detect botnet infections that may send suspicious files to the cybersecurity company for further inspection whenever an issue with this file cannot be resolved locally. The controller must in this case assess whether this functionality complies with the GDPR (e.g., with regard to third country transfer).

4. Concrete scenarios

In this section, we put our general considerations from the previous section into action by illustrating typical scenarios in botnet defence. For every scenario, we focus on particular legal issues arising in this context, while the analysis is by no means exhaustive. The primary aim of these scenarios is to convey a good intuition on the relevant aspects to consider when assessing real-world cases.

4.1. Sinkholing by a public CSIRT

Consider a Computer Security Incident Response Team (CSIRT) located in a security authority (Article 10(1)(2) NIS 2 Directive) requesting an ISP to make the bots' DNS queries resolve to a CSIRT's server IP address (DNS sinkholing). Additionally, the incoming traffic is used for analysis. For our scenario, we assume that the malicious domain has been identified before (e.g., using honeypot data, reverse-engineered malware, traffic analysis). In Germany, the competent authority for cybersecurity (the BSI) performs such activity,[105] legally backed by German legislation.[106] We only consider the processing by the CSIRT.

4.1.1. Applicable data protection regime

With respect to the applicable data protection regime, the question arises whether the GDPR or the Law Enforcement Directive applies, given that the CSIRT in our example is located within a security authority.[107] According to Article 2(2)(d) GDPR and Article 2(1) in conjunction with Article 1(1) Law Enforcement Directive, the latter applies exclusively "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, including the safeguarding against and the prevention of threats to public security." We have to assess, in particular, whether a CSIRT located in a competent authority prevents threats to public safety (i.e., cyberattacks) by disturbing the communication with the C2 servers.

For CSIRTs set up within security authorities (our case), it has been argued that the Law Enforcement Directive spans the applicable data

[98] Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2024:370, para 59.

[99] Similarly, A. Roßnagel and P. Richter, in: I. Spiecker gen. Döhmann et al. (eds.), *General Data Protection Regulation* (Article-by-Article Commentary, Nomos 2023) Art. 5, para 121.

[100] This is not to be confused with the question of whether (trained) machine learning models contain/are personal data; see, e.g., M.R. Leiser and F. Dechesne, 'Governing machine-learning models: challenging the personal data presumption' (2020) 10 International Data Privacy Law, 187.

[101] Note that Article 6(4) GDPR constraints changing the purpose after the initial processing instance.

[102] For a more detailed discussion of the specification criterion and how it relates to other principles, A. Roßnagel and P. Richter, in: I. Spiecker gen. Döhmann et al. (eds.), General Data Protection Regulation (Article-by-Article Commentary, Nomos 2023), Art. 5, paras 53 et seq.

[103] See also in that sense, Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2024:370, para 97.

[104] A. Roßnagel and P. Richter, in: I. Spiecker gen. Döhmann et al. (eds.), *General Data Protection Regulation* (Article-by-Article Commentary, Nomos 2023) Art. 5, para 133.

[105] Bundesamt für die Sicherheit in der Informationstechnik (BSI), 'Die Lage der IT-Sicherheit in Deutschland 2023' <https://www.bsi.bund.de/SharedDocs/Downloads/DE/BSI/Publikationen/Lageberichte/Lagebericht2023.pdf?_blob=publicationFile&v = 7>, p. 14.

[106] Sec. 7c(1), (3) and (4) of the Gesetz über das Bundesamt für die Sicherheit in der Informationstechnik and Sec. 169(5) and (6) of the Telekommunikationsgesetz.

[107] For a broader discussion in the interesting context of public-private partnerships, which may become relevant in cybersecurity contexts, see N. Purtova, 'Between the GDPR and the Police Directive: Navigating through the Maze of Information Sharing in Public-Private Partnerships' (2018) 8 International Data Privacy Law, 52.

protection regime.[108] However, the wording of Article 2(2)(d) GDPR and, more explicitly, Recital 19 GDPR and Recitals 11 and 12 of the Law Enforcement Directive stipulate that only the execution of tasks specifically related to the purposes of the Law Enforcement Directive fall within its scope, hence the institutional setting is not decisive. As to what public security in the sense of this provision means, the CJEU recently ruled that the publication by a public authority of the names of sanctioned traffic offenders does *not* fall within the material scope of the Law Enforcement Directive.[109] This judgment directly affects cases in which a public CSIRT makes available to the public information on currently active botnets (in this respect, Article 23(7) of the NIS 2 Directive). On a more general level, the CJEU implies that the disclosure of names, which aims at ensuring road safety, does not constitute an instance of "preventing threats to public security". Thus, the Court interprets the provision quite restrictively.[110] Now, sinkholing DNS requests, in particular outside the context of public prosecution (recall that since the bots generally do not act criminally there is nothing to prosecute), aims at guarding Internet security, which can be compared to road safety as both aim at maintaining a publicly relevant infrastructure. Based on the reasoning in this case-law, our scenario should be assessed under the GDPR.

4.1.2. Processing of personal data

In our scenario, the CSIRT receives incoming traffic which is meant to be spread across the botnet because the bots' DNS queries resolve to the CSIRT server IP address. Depending on which protocol is used, what the botnet's purposes are and how the bots encrypt the communication, the content received can range from nonsense to a bulk of data stolen by the bot.[111] Typically, however, the bot's IP address will be present. Using the *Breyer* formula from above, the quality of the IP address as personal datum is problematic because the CSIRT may not be able to resolve the IP address itself. Hence it solely depends on whether the CSIRT has (legal) means to gain access to the additional information of the ISP (to which the IP address is related, which is not necessarily the ISP that executes the sinkholing) in order to identify the natural person. In the absence of such means the IP address cannot be classified as personal data and the GDPR would thus not apply to the processing.[112] To be sure, this does not preclude that the content of what the CSIRT receives from the bot can contain (sensitive) information that can amount to personal data, e.g., account data stolen from the infected user containing the user's name.

4.1.3. Legal constraints

The CSIRT, as a public authority, can in principle rely on Article 6(1)(e) GDPR, according to which processing is lawful when necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller (see also, Recital 121 NIS 2 Directive). Therefore, it must be asked what is necessary to achieve the task of ensuring cybersecurity. This requires that the data has been obtained legally, i.e., that the request vice versa the ISP is legally sound.[113] In particular, the legal basis for this request must withstand legal scrutiny under Article 15(1) in conjunction with Article 6(1) and (5) of the ePrivacy Directive, as interpreted by the CJEU.[114] Second, while the processing of the IP addresses is rather unproblematic, in particular since the incoming IP traffic stems from the bot malware and therefore does not disclose any behaviour of the user, a different question arises with regard to sensitive content that might be sent to the CSIRT, if such content is sent by the bot. Such data, in particular when not required for analysis (Article 5(1)(b) GDPR), should be immediately erased, otherwise (as far as possible) made anonymous. In our context, German law additionally sets a maximum limit of three months notwithstanding necessity,[115] although the CJEU does not demand to set a specific time limit to justify the retention of personal data.[116]

4.2. (Network-level) IDS use by a private company

In this scenario, consider an employer using an Intrusion Detection System (IDS) to protect its own network from botnet infections. IDS systems vary widely in what functionalities they offer. We assume that the IDS monitors (and analyses) incoming and outgoing traffic, but also performs Deep Packet Inspection (DPI). While the former analyses the traffic on the level of connection data, DPI inspects the content of the packets sent as far as possible. Assume that our DPI implementation even breaks TLS encryption by redirecting the Internet traffic over a company server (which essentially constitutes a Man-In-The-Middle attack); in such a case, the entire content of the communication is visible. Furthermore, the IDS logs data.

4.2.1. Applicable data protection regime

In employment contexts, i.e., as far as the personal data of employees is concerned, the GDPR applies, while leaving room for Member States' legislation to implement "more specific rules" (Article 88 GDPR). According to Article 88(2) GDPR, however, the main principles of European data protection law must be incorporated into those rules. In pointing to the "relationship of subordination between the employee and the employer" (hence, presumably the particular need to protect the employees' rights) and the GDPR's aim to ensure a high level of protection[117] the CJEU made questionable whether the GDPR's level of protection can be altered 'downwards' by a substantial degree. Here, we assume for simplicity that no national legislation pertains.

4.2.2. Involvement of personal data

Observe that inspecting in- and outgoing traffic to/ from the company network may affect personal data of individuals inside the company network (most importantly, employees) and of outsiders, hence

[108] M. D. Cole and S. Schmitz, 'The interplay between the NIS Directive and the GDPR in a cybersecurity threat landscape' (2020) University of Luxembourg Law (Working Paper No. 2019-017) <https://ssrn.com/abstract=3512093> or <https://doi.org/10.2139/ssrn.3512093>, last accessed 05 February 2024, *infra III.3*: "Thus, if CERTs and CSIRTs are set up within security authorities, the rules on data processing within the Police Directive 2016/680 are applicable, which would require an explicit authorization of CERTs and CSIRTs to process personal data."

[109] Case C-439/19 *B v Latvijas Republikas Saeima* ECLI:EU:C:2021:504, paras 69 et seq.

[110] See also, Case C-817/19 *Ligue des droits humains v conseil des ministres* ECLI:EU:C:2022:491, paras 70 et seq.

[111] B. Stone-Gross and others, 'Your botnet is my botnet: analysis of a botnet takeover' *in Proceedings of the 16th ACM Conference on Computer and Communications Security* (ASM 2009), pp. 635-647.

[112] A subtle issue, which we only note at the side, is the question of whether the CSIRT and the ISP could be considered as joint controllers in the sense of Article 26 GDPR, given that the CJEU states that a person "who exerts influence over the processing of personal data, for his own purposes, and who participates, as a result, in the determination of the purposes and means of that processing, may be regarded as a controller", see Case C-25/17, *Jehovan todistajat - uskonnollinen yhdyskunta* ECLI:EU:C:2018:551, para 68.

[113] In a similar vein, Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2024:370, paras 84 et seq.

[114] For such an interplay between the GDPR and the ePrivacy Directive, Case C-597/19, *M.I.C.M. v Telenet BVBA* ECLI:EU:C:2021:492, para 131.

[115] § 7c(4) of the Gesetz über das Bundesamt für die Sicherheit in der Informationstechnik.

[116] Case C-520/20, *DB, LY v Nachalnik na Rayonno upravlenie Silistra*, ECLI:EU:C:2022:12.

[117] Case C-34/21, *Hauptpersonalrat der Lehrerinnen und Lehrer beim Hessischen Kultusministerium v Minster des Hessischen Kultusministeriums* ECLI:EU:C:2023:270, paras 53-54.

internal and external data. Orthogonally, we can distinguish between traffic data and content data.

As regards the traffic data component (in isolation), the employer takes an internal perspective with respect to the network identifiers of the employees, since, let us say, she assigns local IP addresses to the devices based on login credentials. Hence, local IP addresses are personal data. In turn, the external IP addresses are not resolvable to the employers themselves. Whether means to identify a person exist, must be determined on a case-by-case analysis, as we have seen, taking into account that the employer wishes to protect its network only. It can well be the case that such analysis leads to conclude that external IP addresses are not personal data.

Concerning content data, this data is personal regarding the employees because of its relation to the IP address, which for the employer is personal data. For third-party data, the quality of personal depends on the classification of the traffic data or on what the content is. Here again, the employer must anticipate that even very sensible information (e.g., a very sensitive chat on an online platform) is captured.

4.2.3. Legal determinants

4.2.3.1. Employee data. As a legal basis, the employee could have consented to the processing in the sense of Article 6(1)(a) GDPR. However, it must be ensured that consent is given freely, which requires careful examination in a situation involving social pressure, like the employment context, as indicated in Recital 43 of the GDPR.[118] In that regard, the CJEU specified that even though a dominant position does not rule out that consent is given freely in the sense of Article 4(11) GDPR, this dominant position is an important aspect to be considered within assessment.[119] Among other things, freely given consent requires that the employee is informed accordingly (Article 4(11) GDPR), which might not be sufficiently realised in practice.[120]

In contrast, performance of a contract (Article 6(1)(b) GDPR) does not constitute a suitable legal basis because network security is not "objectively indispensable for a purpose that is integral to the contractual obligation intended for the data subject".[121] Put differently, the employer is in general not *integrally* obliged by the contractual relationship towards the employee to ensure network security, given that there is no direct dependency between the employer's obligation (most importantly, to renumerate the employee) and network security.

If the consent of the employee is not given or not valid, the employer must rely on the legitimate interest clause. Recall that the CJEU in the context of data retention by ISPs for law enforcement purposes when balancing the conflicting interests is concerned, most notably, by the risk of establishing an individual profile.[122] Given the internal perspective of the employer, this risk is realistic in particular when the employee is authorised to use the employer's network for private purposes. To determine in such a case to which extent the employer is allowed to inspect the employee's network data requires taking into account the concrete case. Thereby, it should be borne in mind that, in contrast to the chilling effect on the use of ISP infrastructure (which the CJEU assumes as a consequence of surveillance in the general data retention cases), a chilling effect on the private use of the employer's infrastructure does not seem as problematic. This applies at least to traffic data.

As regards, in particular, content data, the assessment is very complex and requires a broader analysis than can be provided here. For that, note that the European Court of Human Rights (ECtHR) has not categorically ruled out that even content monitoring for the purposes of assessing the behaviour of the employee can be legal under Article 13 of the ECHR. The ECtHR formulated a set of criteria, which includes: whether the subject has been notified, the extent of the monitoring by the employer and the degree of intrusiveness, legitimate reasons for the monitoring, availability of less intrusive alternatives, adequate safeguards for the subject.[123] The judgment, by means of Article 52(3) of the Charter, specifies a *lower* bound on the Charter's fundamental rights; for the GDPR, this becomes relevant when interpreting its provisions in accordance with primary law.[124] Since IDS techniques do not aim at assessing the behaviour of the employee, additional data minimisation efforts can be applied, such as pseudonymization (Article 25(1) GDPR)[125] or decentralized processing: As an example in our context, consider an employee's login credentials tied to the MAC address of her computer (where the MAC address is logged for cybersecurity purposes) and the mapping of the MAC address to the employee's name (stored to keep track of issued devices). Those different pieces of information can remain unrelated for most of the time without affecting each other's purposes. Hence, they should not be linked (and stored in such a manner), since linked data, as we have seen, interferes more with the rights of the data subject.

4.2.3.2. External data. For the external data, the employer can rely on the legitimate interest clause, unless a legal obligation in the sense of Article 6(1)(c) GDPR exists. For traffic data, provided that it constitutes personal data, the legal constraints should not be too tight. In particular, the risk of establishing a profile is, as for external traffic data in general, typically rather low. Whether capturing content data is necessary and proportionate with respect to the employer's legitimate interest in ensuring cybersecurity, is more difficult to answer. Given that the third-party user typically expects (encrypted) communication to remain secret, in particular since the user cannot reasonably detect the interception, it seems hard to justify at least logging such data (as opposed to a transitory scanning).

4.3. Honeypot use by an ISP

As we have already noted, (large) ISPs are in an auspicious position to combat botnets because they have, in principle, access to accumulated data from multiple (private) networks and can associate IP addresses with access points (served by them).[126] Let us here consider a scenario in which an ISP deploys a honeypot to attract botnet traffic. A peer-to-peer botnet has fallen for the honeypot and communicates the IP addresses of the adjacent bots to the honeypot. As we have seen, the scenario must be

[118] EDPS, 'Guidelines 05/2020 on consent under Regulation 2016/679' (2020) <https://edpb.europa.eu/sites/default /files/files/file1/edpb_guidelines_202005_consent_en.pdf> pp. 13 et seq last accessed 25 October 2024. See also, J. Rauhofer and B. Schafer, in: I. Spiecker gen. Döhmann et al. (eds.), General Data Protection Regulation. (Article-by-Article Commentary, Nomos 2023) Art. 4(11), para 25.

[119] Case C-252/21, *Meta v Bundeskartellamt* ECLI:EU:C:2023:537, paras 147-148.

[120] According to a dataset from Canada and the US—where the GDPR does not apply—few employees consider themselves well informed about what data the employer collects for cybersecurity purposes, see J. Stegman and others, '"My privacy for their security": Employees' privacy perspectives and expectations when using enterprise security software' (2022), *arXiv preprint arXiv: 2209.11878*. Recall that employment contexts are subject to possible derogations by the Member States. However, transparency of processing must still be ensured according to Article 88(2) GDPR.

[121] Case C-252/21 *Meta v Bundeskartellamt* ECLI:EU:C:2023:537, para 98.

[122] See *infra* 3.2.3.

[123] *Bârbulescu v. Romani*, App no 61496/08 (ECtHR, 05 September 2017) ECLI:CE:ECHR:2017:0905JUD006149608, para 121.

[124] See, as regards such interpretation in general, e.g., Case C-583/13 P *Hoechst v Commission,* ECLI:EU:C:1989:337, para 12.

[125] F. Menges and other, 'GDPR-compliant data processing in modern SIEM systems. (2021) 103 (102165) Computers & Security 1.

[126] H. Asghari, M. J. Van Eeten, and J. M. Bauer, 'Economics of fighting botnets: Lessons from a decade of mitigation' (2015) 13 IEEE Security & Privacy 16.

evaluated primarily according to the ePrivacy Directive.

4.3.1. Processing of personal data

In a honeypot case, the potentially personal data processed are mainly the IP addresses of other members of the botnet. Recall that we generally assume that IP addresses are personal data from the perspective of an ISP, with the constraint that this only applies to IP addresses assigned by this ISP. Thus, only with respect to those the classification of IP addresses as personal data is given by assumption. Assume that there are at least some of those contained in the honeypot data. For other IP addresses, this again depends on whether reasonable means to resolve the IP address exist.

4.3.2. Legal basis

It has been questioned in the literature whether the processing of connection data by ISPs after service delivery is lawful if this data is processed in a non-anonymous (i.e., personal) form and there is no consent from the subject concerned.[127] Moreover, the authors of this concern believe that the collection of IP addresses is typically mandatory under national law and acknowledge that the legal framework provides for exceptions that cover such national legislation; however, simply storing IP addresses without deriving further information does not reveal anything about the prevalence of botnets (which is obvious).

However, for ISPs there is a legal basis for deploying botnet tracking techniques such as the use of honeypots. Article 15 of the ePrivacy Directive allows national legislation to override the Directive's default rule according to which traffic data should, in principle, be erased/anonymised after the communication instance (Article 6(1) ePrivacy Directive). One such exception concerns provisions aimed at the "prevention, investigation, detection and prosecution of criminal offences or of unauthorised use of the electronic communication system", which, given that botnet activity constitutes criminal behaviour under European law, comprises botnet defence. In this sense, the CJEU explicitly linked the unauthorised use of the communication system to the integrity and security of the network.[128][129] Furthermore, Article 4(1) ePrivacy Directive contains an obligation of ISPs to ensure the security of their networks; Article 4(1a) defines a set of topics that shall be implemented "without prejudice to Directive 95/46/EC" (Data Protection Directive). Therefore, upon transposition by the Member States, the ePrivacy Directive itself requires the ISPs to introduce cybersecurity measures, acknowledging that those might impact data protection (and require balancing).

4.3.3. Legal constraints

When it comes to ISPs and Internet traffic monitoring, there is a particular concern that such activities can lead to the establishment of profiles of individuals, at least in the context of general and indiscriminate data retention for law enforcement purposes. Intuitively, this results from the role of ISPs as gatekeepers in granting access to the Internet and thus having the possibility to map IP addresses onto access points and potentially analysing traffic at a large scale.[130] This is again where the chilling effect comes in Recall that the CJEU's jurisprudence regarding indiscriminate data retention concerns the obligation of ISPs to retain data of its customers.

But in contrast to general data retention for the purposes of identifying individuals by law enforcement, processing personal data for the purposes of cybersecurity differs substantially in terms of privacy intrusiveness. First, cybersecurity measures are, in principle, not dependent on the identification of individuals, while the collection of data for law enforcement purposes precisely aims at identifying a (suspect) individual. Second, any identification, even if required (e.g., to notify the individual), would not take place in the context of law enforcement activities by the state, in the course of which safeguarding fundamental rights is of particular importance.[131] In addition, network security lies in the interest of the customers, which is highlighted by the ISPs' duty to inform its customers under certain circumstances of a data breach (Article 4(3) ePrivacy Directive).

Furthermore, in our scenario, a honeypot is applied, which can be seen as a targeted method in that it processes data (typically) related to a botnet (as opposed to indiscriminate monitoring). In contrast to general and indiscriminate data retention, the individuals concerned are delimited by their participation in the botnet. Furthermore, such data does not represent individual actions of the infected individual since the action attracted by the honeypot is controlled by the bot malware; from this point of view, the risk of establishing an *individual* profile is low. Overall, therefore, there are very significant differences with the CJEU's case law regarding generalised data retention. Finally, it should be recalled that according to recent case law the impact on the data subject can be alleviated by implementing a strict organizational separation between the data and the individualizing information.

4.4. Information exchange between private companies

It is a generally accepted fact that information sharing between individual networks holds promise for improving cybersecurity performance.[132] This is particularly true for botnets, given their decentralized structure. There are already public and private attempts to foster sharing; for instance, the public CSIRT in Luxemburg operates an easily accessible server for the exchange of cybersecurity information, based on the open-source MISP platform infrastructure.[133] For our scenario, consider a private company that registers with this platform and contributes information.

4.4.1. Involvement of personal data

The MISP data format for information exchange[134] contains fields that might contain personal data.[135] This could be again IP addresses[136] marked, e.g., as bot-related. Recall that at least according to the General Court, the average perspective of the receiving entities is decisive in determining the classification of the IP address as personal data for the purposes of the disclosure of this address.[137] Hence, it is a plausible consequence that an IP address is generally not a personal datum in this

[127] L. Böck and others, 'Processing of botnet tracking data under the GDPR' (2022) 45 (105652) Computer Law & Security Review 1, 16.

[128] Case C-275/06 *Promusicae v Telefónica des España SAU* ECLI:EU:C:2008:54, para 52.

[129] Note that the mandate to introduce such provisions is not limited to using specific data or types of processing, especially not to the mere collection of IP addresses (note also the explicit use of "inter alia" in the wording of Article15(1) of the Directive).

[130] A. Martin, &, N. N. G. de Andrade Battling, 'Botnets with Digital Rights in Mind' (2012) 3 (2) European Journal of Law and Technology 1.

[131] Analogously, Case C-470/21 *La Quadrature du Net v Premier ministre* ECLI:EU:C:2022:838, Opinion of AG Szpunar, paras 55 et seq.

[132] F. Skopik, G. Settanni, and R. Fiedler, 'A problem shared is a problem halved: A survey on the dimensions of collective cyber defense through security information sharing' (2016) 60 Computers & Security 154.

[133] <https://www.circl.lu/services/misp-malware-information-sharing-platform/>.

[134] <https://www.misp-standard.org/rfc/misp-standard-taxonomy-format.html>.

[135] For a general overview of the disclosure risks in, among others, the MISP data structure, V. Jesus., B. Bains, & V. Chang, 'Sharing Is Caring: Hurdles and Prospects of Open, Crowd-Sourced Cyber Threat Intelligence' (2023) 71 IEEE Transactions on Engineering Management 6854.

[136] M. Horák, V. Stupka, & M. Husák, 'GDPR compliance in cybersecurity software: A case study of DPIA in information sharing platform' *in Ares '19: Proceedings of the 14th international conference on availability, reliability and security,* (ASM 2019).

[137] See the discussion *infra* 3.1.1.

context. Furthermore, information exchange on the strategical or tactical level typically does not contain personal data. But it is possible that some personal data is involved, e.g., a domain name (in particular cases).[138]

4.4.2. Legal basis

Regarding the availability of a legal basis for the processing, the voluntary exchange of information can be, in principle, justified by Article 6(1)(f) GDPR (which corresponds to the assessment in Recital 121 NIS 2 Directive). In this context, an important normative goal of the EU legislator is the aim to actively foster information exchange in cybersecurity (Recital 120 NIS 2 Directive). In particular, Member States must, according to Article 29 NIS 2 Directive, "ensure that entities falling within the scope of this Directive and, where relevant, other entities not falling within the scope of this Directive are able to exchange on a voluntary basis relevant cybersecurity information", where "[s]uch exchange shall be implemented through cybersecurity information-sharing arrangements in respect of the potentially sensitive nature of the information shared."[139] With this, the legislator wishes to tackle the reluctance to share information due to, among other things, legal uncertainties perceived by the entities (Recital 119).[140] Hence, and having particular regard of that, the interest of sharing must be weighed high when performed within an information arrangement crafted with rules concerning data protection.

Given the transnational span of botnets, it is desirable to enable sharing with entities from third countries. However, the transfer of personal data is subject to additional requirements pursuant to Article 44 GDPR. The fulfilment of these additional requirements can be quite strenuous, which has been highlighted by the CJEU's judgments in *Schrems I* and *Schrems II*.[141] The NIS 2 Directive explicitly aims to allow CSIRTs to exchange relevant information with third countries' CSIRTs, including personal data in accordance with the GDPR (Article 10(7) NIS 2 Directive). Since cybersecurity, as we have mentioned before, constitutes a public interest, information exchange by CSIRTs can fulfil the respective requirement in Article 49(d) GDPR, which demands for such transfers an important reason of public interest to pertain. For (private) entities in general, however, it is at first glance questionable whether such an important reason of public interest is given. In the absence of an adequacy decision, the solution may lie within the use of standard contractual measures (within information-sharing arrangements in the sense of Article 29 NIS 2 Directive) in conjunction with technical measures that qualify as effective supplementary measures in the sense of the *Schrems II* judgment of the CJEU.[142] For instance, Secure Multiparty Computation Techniques[143] may be used to restrict information sharing to certain thresholds[144] or to derived information such as the relation between cybersecurity events without disclosing the data responsible for establishing this relationship.[145] For machine learning approaches, federated learning is a promising concept, which in research is being explored in the context of cybersecurity and, in particular, the MISP infrastructure.[146]

5. Conclusion

As the digital landscape evolves, so too does the complexity of threats such as botnets, necessitating advanced defence mechanisms. However, in the European Union, the implementation of these mechanisms must navigate the requirements of data protection laws, notably the GDPR. This paper has highlighted the tension between the necessity for robust cybersecurity measures and the imperative to protect individual privacy rights under EU law.

Most importantly, we observe that the current discussion underestimates the implications that the interplay of different legal regimes, the unclear interpretation of fundamental concepts of the GDPR (such as the concept of identifiability with particular regard to IP addresses), and the jurisprudence of the CJEU has in assessing this tension. In constructing four different scenarios, we have illustrated what implications this entails for the complexity of the legal assessment of a particular case. For instance, it is yet unclear in which context and to which extent the CJEU's (restrictive) jurisprudence concerning indiscriminate data retention for law enforcement purposes affects the assessment of cybersecurity techniques. Therefore, there is a need to develop a nuanced discussion for the particular scenarios, taking into account the legal issues raised here.

Given the importance that the EU legislator attaches to ensuring cybersecurity, we should be diligent in applying data protection rules not to discourage the effectiveness of the respective endeavours in practice. However, a more careful analysis is required whenever the controller has, potentially, the possibility to (mis)use the data to establish a profile of the data subject. In such cases, implementing technical and organisational procedures that ensure that the individual's data is processed separately holds promise to sufficiently alleviate these risks.

[138] <https://www.misp-project.org/compliance/GDPR/>.

[139] The current draft for the transposition of the NIS 2 Directive into German law provides that the BSI shall operate an online sharing platform, see sec. 6 of the Referentenentwurf für das Gesetz zur Umsetzung der NIS-2-Richtlinie und zur Regelung wesentlicher Grundzüge des Informationssicherheitsmanagements in der Bundesverwaltung, available at <https://ag.kritis.info/wp-content/uploads/2023/07/NIS2UmsuCG-Referentenentwurf-BMI-CI1-Bearbeitungsstand-03072023.pdf>.

[140] For a literature review, A. Zibak, A. Simpson, 'Cyber Threat Information Sharing: Perceived Benefits and Barriers' *in Ares '19: Proceedings of the 14th international conference on availability, reliability and security* (ASM 2019).

[141] Case C-362/14 *Schrems v Data Protection Commissioner* ECLI:EU:C:2015:650; Case C-311/18, *Data Protection Commissioner v Schrems* ECLI:EU:C:2019:1145.

[142] Case C-311/18 *Data Protection Commissioner v Schrems* ECLI:EU:C:2019:1145, para 133.

[143] Secure Multiparty Computation refers to cryptographic protocols that enable two or more parties to jointly compute a function without disclosing the input data. Every party only receives the common output. The respective techniques are also mentioned in EDPB, 'Recommendations 01/2020 on measures that supplement transfer tools to ensure compliance with the EU level of protection of personal data" (2021), para 138. For a GDPR-based analysis, see L. Helminger and C. Rechberger, 'Multi-party computation in the GDPR' in *Privacy Symposium 2022: Data Protection Law International Convergence and Compliance with Innovative Technologies (DPLICIT)* (Springer 2022), pp. 21–39; A. Treiber, D. Müllmann, T. Schneider, and I. Spiecker genannt Döhmann, 'Data protection law and multi-party computation: Applications to information exchange between law enforcement agencies' in *Proceedings of the 21st Workshop on Privacy in the Electronic Society* (2022), pp. 69–82 (the latter in the context of information exchange between legal agencies).

[144] R. A. Mahdavi and others, 'Practical over-threshold multi-party private setintersection' in *Annual Computer Security Applications Conference*, (ACM 2020) pp. 772–783.

[145] Davy Preuveneers and Wouter Joosen, 'Privacy-preserving correlation of cross-organizational cyber threat intelligence with private graph intersections' (2023) 135 (103505) Computers & Security 1 (with a MISP implementation).

[146] J. R Trocoso-Pastoriza and others, 'Orchestrating collaborative cybersecurity: a secure framework for distributed privacy-preserving threat intelligence sharing' (2022) arXiv preprint arXiv:2209.02676.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was supported by a Research Grant from the Deutsche Forschungsgemeinschaft (DFG) as part of the project "Lawful Detection, Investigation, and Prosecution of Botnet-related Crime" (project reference: 457175502, record number: SO 1133/1–1). Many thanks to Leo Dessani, Max Leicht, Christoph Sorge, and—last, but definitely not least—Nils Wiedemann for their highly valuable comments as well as to Niklas Dahlem for helping with the final formatting. Furthermore, I benefited very much from the comments of an anonymous reviewer.

Data availability

No data was used for the research described in the article.