Nothing to hide? On the Security and Privacy Threats beyond Open Data

Javier Pastor-Galindo, Félix Gómez Mármol, Gregorio Martínez Pérez
Department of Information and Communications Engineering
University of Murcia, Spain
{javierpg, felixgm, gregorio}@um.es

***Abstract*—People's online activity continuously dumps personally identifiable information on the web. Alarmingly, this public information becomes a dangerous cyberweapon in the era of finely-targeted cyberattacks. This article explores today's cyberthreats around open data, from traditional ones to AI empowering, thus unveiling a range of vulnerabilities to which end-users are exposed.**

Index Terms: Security, Abuse and crime involving computers, Unauthorized access (hacking, phreaking), Internet security policies.

1. A globally connected world

Interconnected applications, social media services, and online platforms are booming. Users spend a disproportionate amount of their time in these virtual places, leading to an overexposed lifestyle [1] and flooding the open web with personally identifiable information (PII) [2].

This context enables high Open Source Intelligence (OSINT) opportunities with a significant social impact [3]. However, the lack of privacy in online behaviour also becomes a delicate vulnerability to be exploited by cybercriminals for obtaining sensitive information and designing tailored cyberattacks [4].

Unfortunately, those privacy-compromising cyberattacks have a high success rate [2]. Considering that recently reported attacks even present signs of Artificial Intelligence (AI) weaponization [5], a tricky scenario for end-users emerges.

In this line, researchers should identify which are those cyberattacks where cybercriminals employ open data, i.e., public information obtained in a completely lawful manner from sources such as search engines, social networks, or websites, without the need for special permissions [3]. Consequently, it is possible to report privacy vulnerabilities and stimulate the research of such cyberthreats and associated countermeasures.

However, to the best of our knowledge, there is a lack of research works addressing cyberattacks with open data, which is particularly important to link end-users exposure with its actual cyberrisk. To fill the gap, the contributions of this article are (a) the compilation of a wide variety of traditional and brand-new cyberattacks specifically launched or complemented with publicly available information, (b) exploration of the latest tendencies in AI-powered offences, particularly focused on cyberattacks with open data, (c) proposal of a taxonomy to classify data-based

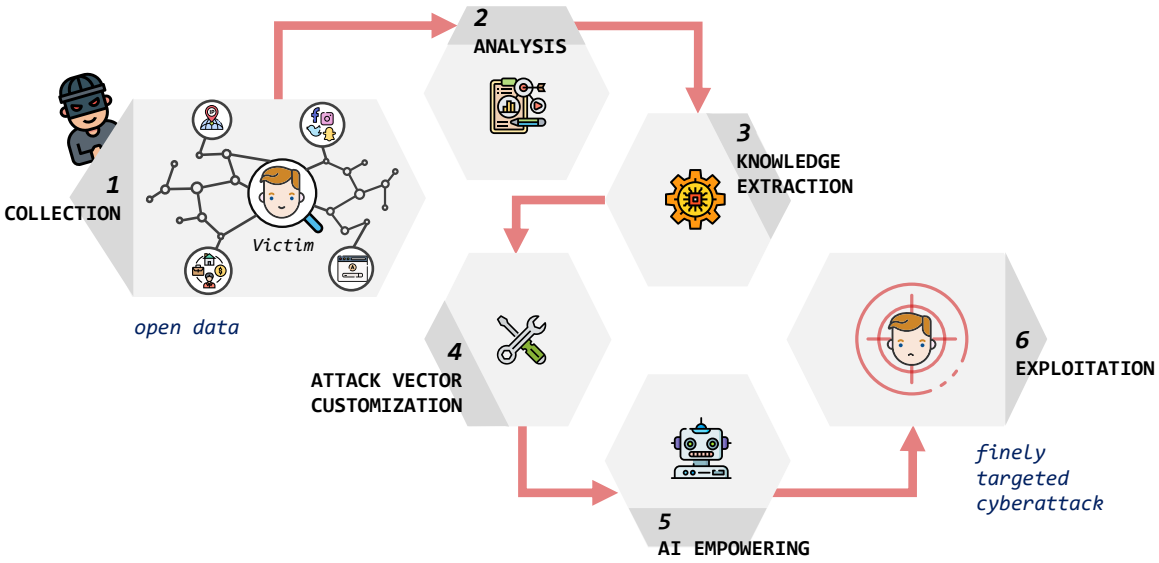

Figure 1. Construction of a finely targeted cyberattack with OSINT

cyberattacks depending on how they use such open data, and (d) suggestion of practical habits for end-users to minimize risks, and enumeration of principles to be adopted by companies and academia/industry to contribute to world-wide users security and privacy.

2. From open data to a successful cyberattack

The latest *Sophos Threat Report* [6] or the *CheckPoint Security Report* [7] reveal the criminal tendency of profiling the target to fuel more accurate offences.

As shown in Figure 1, leveraging OSINT, the cybercriminal collects open data about the victim, analyzes it to extract PII (e.g., real names, job positions, email addresses, telephone numbers, social media information, company data, public files, etc.), and finally infers advanced knowledge (economic situation, preferences, habits, personal beliefs, activity on the web, etc.) [3]. With that intelligence, the cybercriminal tailors the appropriate attack vector that exploits specific weaknesses. Furthermore, the employment of AI has complemented and fortified traditional threats, thus achieving higher effectiveness and even generating brand-new attack vectors [8]. AI-based cyberoffence is able to learn from successful intrusions, deceive algorithms, and evade robust defence systems [4]. Considering that AI models are usually based on data-driven approaches, open data becomes a dangerously exploited goldmine.

Several OSINT-based attacks are reported daily. In December 2019, end-users using the genetic *GEDmatch* service had their genotypes revealed on the web. In July 2020, *Twitter* suffered a cryptocurrency scam that affected 130 high-profile accounts through a phone spear-phishing attack to some employees, presumably fueled with pretexts taken from open sources. Also, in July of the same year, some *Garmin* employees were profiled to discover their visited websites. The latter were infected with fake software that triggered a ransomware attack that disabled *Garmin* services for three days. Some other cyberthreats are explicitly using AI techniques: social bots that profile victims in political campaigns on social media, the production of tailored fake news, the spread of never seen malware, or massive controlled DDoS attacks, among others [5], [9].

3. Connecting the dots with OSINT: Attacks through public information

The latest technical advances in cyberoffence permit the cyberattacks to have finely-targeted reconnaissance, artificial replication of human capabilities, undetectable and adaptable malware payloads, automated behavior, and scalable scope [4], [5], [8]. All this, in conjunction with the massive

volume of public information available, provokes serious cyberrisks.

In Table 1, we classify the most representative cyberattacks with open data and their AI-based evolution, obtained in our review. In this sense, we suggest a taxonomy to categorize them depending on how they benefit from used data since, as far as we know, there is no classification in the literature following that perspective [10].

Firstly, we identify three threat dimensions that determine how the cybercriminal uses open data to harm the victim. In the threat of *deception*, data is employed to mislead victims or systems. In the *blackmail* threat, the information is applied to pressure or manipulate the victims. Finally, in the *expansion* threat, data is the means to acquire more sensitive information and extend the potential attack surface.

Secondly, each threat dimension is subdivided into different types of attack, which particularly reflect the methodology. In the *deception* threat, open data is utilized to cheat victims by practicing *social engineering*, *spear-phishing* or *impersonating* another person, and bypass systems through unauthorized *intrusion*. Regarding the threat of *blackmail*, open data is strategically applied to directly manipulate the target through *extortion*, *defamation*, or *persuasion*. Within the *expansion* threat, open data fuels knowledge generation processes of *monitorization*, *profiling*, and *recognition*.

Note that some cyberattacks may have different intentions, so they fit into various categories. Complementarily, we present, for each type of attack, a list with the public traces (enablers) that facilitate them. All results are in-depth explained by threat dimension in Sections 4, 5, and 6.

4. Threat of deception

The first type of threat, deception, implies the strategic manipulation of certain information to deduce sensitive details, fraud victims with tricks, or gain access by circumventing authentication systems.

4.1. Social engineering attack

The most common cyberthreats based on social engineering are the so-called *phishing attacks*. Their main objective is to obtain sensitive information from a random group of persons to achieve the offence [11]. Considering that a generic email, a suspect call, or a non-contextualized conversation may have a low success due to a lack of personalization [12], the tailored contact is a more effective manner of enticing the victims [13]. Additionally, AI is employed to learn from successful phishing attacks and bypass detection systems with supervised techniques or adversarial approaches [8].

A *watering hole attack* [14] or a *fake software attack* [15], in which the cyberattacker hacks the victim by compromising or spoofing a website that he/she frequently visits, would be highly effective if the victim has been profiled previously with OSINT. In the same direction, the *OAuth attack* [7] is being used to gain full access to emails by deploying fraudulent delegated authorization services in malicious pages tailored to certain persons. Analogously, the *road apples* or *baiting attack* [15] based on a succulent and malicious URL for the victims through email, or the *tailgating/piggybacking attack* [15] consisting of accessing forbidden spaces, could be perilously built with existing public details. These scenarios become even more dangerous when AI adversarial training is employed to skip malware detection mechanisms [8].

Moreover, specific data required for the cyberattacker could be extracted through premeditated *reverse social engineering attacks* or *pretexting attacks* [15].

4.2. Spear-phishing attack

Target-specific phishing against a selected victim or group is known as a *spear-phishing attack*, which is more likely to succeed than a rudimentary phishing attack [12]. traditionally launched via email. Other frequently published details, such as the telephone number, social network profiles, and usernames, can also be exploited in *smishing attacks* [15]. Victims are deceived via SMS, social networks, or gaming platforms, where the contact is simple, and the number of potential victims is huge [7]. In this sense, bots automate such offence subjugating the victims with tailored messages [11].

The victim could also be manipulated through a voice call in *vishing attacks* [6]. Although the cybercriminal was easy to discover in the past, modern Deep Neural Networks can now replicate

Threat dimension	Type of attack	Representative cyberattacks	AI empowering	Main enablers
Deception	Social engineering	Watering hole attack, Fake software attack, OAuth attack, Baiting attack, Tailgating/piggybacking attack, Reverse social engineering attack, Pretexting attack, Spear-phishing attack	AI-based phishing generation, AI-based domain generation	Email address, location, tastes, preferences, routines, location, economic situation
	Spear-phishing	Representative cyberattacks of Social engineering type, Ransomware attacks, Disaster fraud, Online dating/romance scam, Surrogacy/Adoption scam, Rentals housing scam, Sports Memorabilia fraud, Pet adoption scam, 419 Nigerian scam, Stuxnet attack, Physical-based attack, Clickjacking attack, Whaling attack, Business email compromise attack, Vishing attack, Smishing attack, Robocall attack	AI-based phishing generation, AI-based domain generation, Automatic tailored content generation, Mass delivery, Social bots, Deepfakes, Wetware attacks	Email address, location, tastes, preferences, routines, location, economic situation, real name, username, birth date, job position, telephone number, home address, photos, videos, social network profile, relationships
	Impersonation	Email-based phishing attack, Vishing attack, Smishing attack, Spam attack, House stealing or deed fraud, Identity cloning attack, Cross-profile attacks, Print attack, Replay attack, 3D mask attack, Fingerprint attack	Identity-cloning social bots, AI-supported voice synthesis technologies, Biometric attacks, Deepfake-enabled video attacks	Real name, email address, telephone number, friendships, social network profile, identification numbers, photos, videos, personality, behavior
	Intrusion	Print attack, Replay attack, 3D mask attack, Fingerprint attack, Brute force attacks	AI-supported voice synthesis technologies, Biometric attacks, Deepfake-enabled video attacks, AI-based password brute force attacks	Email address, filtered passwords, real name, user name, photos, videos, network information, preferences
Blackmail	Extortion	Email-based phishing attack, Vishing attack, Smishing attack, Spamming attack, Identity cloning attack, Cross-profile attacks, Grandparent scam	AI-supported voice synthesis technologies, Deepfake-enabled video attacks	Real name, email address, friendships, family ties, social network profile, sensitive information
	Defamation	Smear campaigns, Doxxing attack, Manipulated media, Disinformation messages, Fake news	Deepfake-enabled video attacks, Social bots	Historical facts, social network profile, forum posts, public opinions and attitudes
	Persuasion	Wetware attacks, Natural language generation techniques	Spamming attack, Context-ware spam attack, Sybil attack	Email address, user name, social network profile, online activity
Expansion	Monitorization	Corporate espionage, Sybil attacks, Advanced Persistent Threat	Social bots	User name, social network profile, online activity
	Profiling	Cyberdiscrimination, Inference attack, Illation attack, Graph-based attack, Neighborhood attack, Deanonymization attack, Sybil attack	Social bots	Main enablers of Spear-phishing type, sensitive opinions, network information
	Recognition	Port scanning, Fingerprinting, Footprinting, Vulnerability discovery, Devices reconnaissance, Traditional network-based cyberattacks	Automation of vulnerability discovery, AI hacking tools, AI network-based cyberattacks	Domain, email address, IP address, hostname, DNS records, protocols metadata

Table 1. Threats, types of attacks and representative cyberattacks together with the improvements introduced by AI and the exploited open data

.

the voice of trusted individuals [6]. Particularly, in the *robocall attack* [15], the cybercriminal trains a voice-based conversational agent to impersonate a remarkable figure and propagate the voice message to the list of victims. Training recordings and phone numbers could be crawled from open sources. Furthermore, this AI-based technique is even capable of tricking speaker recognition mechanisms [6].

The *spear-phishing attack* that is generated explicitly towards senior executives of companies is also considered a *whaling attack* [2]. The *business email compromise (BEC) attack* [15] is a representative example in which the social engineer seeks confidential information to jeopardize the company.

It is worth highlighting that no large amounts of data or complex inferences are especially

needed in a spear-phishing attack: the location of users make them susceptible to *disaster fraud* in which the cybercriminal deceives with the excuse of a recent catastrophe; a post with the marital status or a public dating profile may be utilized to perpetuate an *online dating/romance scam* or *sweetheart swindle*, and convey a false sense of love to end up asking for money; the name of a best friend, extracted from social networks, can be used as the sender of a message that asks for a favor or sensitive information; a public message expressing the impossibility to bear, or the unwittingly misuse maternity forums, make the person susceptible to suffer *surrogacy* or *adoption scams*; the publication looking for a house, or an existing account in online floor sales platforms, denotes a perfect profile to be targeted in a *rentals housing scam*; the public fanaticism of users for sports or animals increments the probability of suffering *sports memorabilia frauds* or *pet adoption scams*; the home address linked to a public photo, or the job place on a exposed resume, may facilitate a *Stuxnet attack* [4] by dropping a malicious USB, a *physical-based attack* [15] to steal confidential information, or even a traditional robbery; any trace of personality can fuel the *419 Nigerian scam*, in which cybercriminals intentionally adapt advanced fee techniques to individual circumstances, or a *clickjacking attack* [11], in which messages or posts are tailored with personal tastes to augment the probability of clicking in malicious URLs [4].

The incidences of spear-phishing are on the rise, turning in the worst cases into *ransomware attacks* [6]. Additionally, the recent *wetware attacks* [12] are combining machine-learning-generated content, personification, and evasion to launch massive spear-phishing campaigns [8].

4.3. Impersonation attack

As seen before in *vishing*, *smishing*, or *email-based attacks*, open data is enough for perpetrating unauthorized activities under a forged identity [12].

Simple cyberthreats could include the insertion of obtained phone numbers in online forums or dating sites, the introduction of collected email addresses in spam sites to carry out *spam attacks* [11], or the *house stealing/deed fraud* in which identification numbers, frequently exposed in government documents, are maliciously used to activate transactions.

In sophisticated threats, a cybercriminal could create a social network profile with plagiarized public information of an existing person in an *identity cloning attack*. If the impersonated victim already has a valid account [12], the *cross-profile cloning attack* alternatively copies the identity of the victim from one social network to another where the target does not have a profile [14]. Therefore, the absence of a profile in a service could be considered as a vulnerability [13]. Moreover, AI scales up those threats with *identity-cloning social bots* [12] that implement Natural Language Understanding and Natural Language Generation to interact as humans.

On the other hand, it is currently feasible to imitate an original speaker voice with Neural Networks [8]. Unfortunately, this may derive in *biometric attacks* trained with collected audio/video records to reproduce the authorized voice and bypass authentication systems [12].

Finally, a similar scenario applies to face recognition systems, which could be dodged by AI-powered offences, such as *deepfake-enabled video attacks* [6]. In particular, the employed Neural Networks could be trained with public videos and pictures of authorized individuals to gain access. However, there are other not-so-sophisticated threats affecting recognition systems, such as the cybercriminal presenting collected photos (*print attack* and *fingerprint attack*), videos (*replay attack*), or even a three-dimensional artificial head (*3D mask attack*).

4.4. Intrusion attack

The fight against unauthorized access persists in the cybersecurity field. Traditional intrusions evolve due to AI and data mining techniques, particularly acute with the massive data breaches continuously published in the surface or deep web.

This type of cyberattack is frequently related to impersonation attacks, in which the offender poses as a legitimate user to gain access through biometric or password authorization systems. Apart from previously discussed biometric weaknesses, the password-based schemes are also vulnerable. In such scenarios, the most straightforward way of intruding would be to

search hacked credentials of the victim on open "paste sites" or the deep web. Another possibility would be to reset passwords guessing the typical security questions with personal facts extracted from social networks.

Regarding the *brute force attacks* based on a dictionary, AI improves password cracking in the so-called *AI-based password brute-force attacks* [5]. The latter could employ Generative Adversarial Networks to guess passwords with much higher accuracy or execute Recurrent Neural Networks to recognize patterns in old passwords to achieve a higher success rate [8].

5. Threat of blackmail

In the second type of threat, blackmail, the cybercriminal employs the obtained data to pressure victims, create false sensations, and force them to consciously develop determined actions.

5.1. Extortion attack

Some impersonation cyberattacks may be motivated to conduct an extortion attack. The evildoer poses as a trusted contact to slyly compel the victim to complete a delicate task.

On the other hand, managing certain information, such as tastes, intimacies, weaknesses, family ties, or social circle, by the adversary enables severe psychological manipulations and blackmailing. The possession of these details could demonstrate a purported authority and extortionate the target [14]. The *grandparent scam* is a representative example of extortion where a ransom is required to solve a supposed problem from a relative.

5.2. Defamation attack

People change their way of thinking and behaving over time. Somehow, this history persists within different public sites, such as social networks, media platforms, online forums, or search engines cache. Over time, the outdated content may be unearthed to spoil the integrity of the victim.

The leakage of certain information compromising a company or the public attitude of an employee can motivate the crowd or competitors to trigger a smear campaign against the corporation. Public videos starring a public authority or celebrity could be manipulated by strategically cutting out specific messages, mixing scenes, and unifying a storyline that recreates a situation that never happened. The resulting fake video with decontextualized fragments could expose the victim supporting an unconstitutional act or making a despicable statement. In general, anyone could harm another person with a deliberate selection and publication of sensitive details in the so-called *doxxing attacks*.

Public information can also be manipulated to harm the owners. Disinformation campaigns or fake news are recent examples, which grow in credibility when appropriately accompanied by real facts. Moreover, with the rise of deepfakes, artifacts such as public videos, images, and recordings can train AI face-swapping algorithms to replicate victims. In this way, it is possible to artificially recreate illegitimate scenes or embarrassing speeches that never happened [8]. And to top it off, bots can quickly and massively spread all that malicious content [9].

5.3. Persuasion attack

A malignant persuader will select a group of like-minded victims to match specific properties. The most common activities would include political actions, advertisement spam, or product information flooding [2]. Moreover, the employment of fine-granularity information would build smaller clusters with more eminent tailored propaganda [4].

In persuasion attacks, the *sybil attack* [11] is also employed to create fake allies to spread a sense of legitimacy and make the persuasion more effective. These malicious nodes could be bots that artificially manipulate micro-environments through segmented spam, fake ads, and collusive publicity [12] or social bots that politically polarize society [9]. With the colossal activity registered in media platforms, the interference of malicious actors is raising the cases of automated persuasion campaigns [4], which scale even further if the organic content is automatically created in *wetware attacks* with Natural Language Generation techniques [6].

There are also unsophisticated persuasion activities such as the usage of private telephone numbers in registration procedures, the introduction of email addresses within malware and advertising diffusion repositories, or the flooding of on-

line profiles [11]. Concretely, the owners of email addresses could fall into deceptive advertising networks and suffer *spamming attacks* [2]. These messages mainly contain fraudulent offers and malware, and they could be propagated through victim contacts. Specifically, *context-aware spam attacks* are considered to be more productive and authentic due to personal information in their corpus [14].

6. Threat of expansion

In the third type of threat, expansion, cybercriminals benefit from open data to infer even more considerable knowledge about victim's details, intimacies, and surroundings.

6.1. Monitorization attack

A monitorization attack aims to collect as much information as possible about the online activity of the target. The cybercriminal could also wait for specific abnormalities in the inferred patterns to launch malicious actions, such as when the target has gone away on a trip, has suffered a tragic loss, or has uncovered some secret.

This type of attack reaches a supreme scale in *sybil attacks* [11] thanks to social bots that automatically record the activity of users and detect anomalies in events.

Therefore, being tracked is treacherous, as any mistake or irresponsibility may dramatically benefit the offender. If the monitoring remains over time, it could be considered an *advanced persistent threat* (APT) [14]. *Corporate espionage* [2] is a representative example of continuous observation for commercial purposes.

6.2. Profiling attack

The profiling attack painstakingly studies the target to discover intimate facts such as the economic, political, medical, or religious status. This type of attack is usually helpful for cybercriminals in identifying potential targets [4] or initiating the reconnaissance phase of the Cyber Kill Chain framework [5].

Purchasing interests could be inferred from comments and ratings of public recommender systems. With online behavior, it could be possible to estimate the wealth or willingness to acquire a specific product [4]. Political preferences could be deduced from social network profiles to segment society by ideology and, consequently, organize targeted election campaigns [4] or direct social bots [5]. Another fatal consequence of profiling attacks is cyberdiscrimination when finding a job or negotiating insurance [11].

The social environment of the target is also precious for this type of attack, being especially exploited in *inference attacks* [11]. The so-called *illation attacks* [2] even incorporate network analysis to guess unexposed details, and a *graph-based attack* [2] analyzes the existing contacts and relationships to infer sensitive details. Particularly, a *neighborhood attack* would pursue the location of the victim in the social graph. From that point, it is feasible to detect identities across several social networks [13].

Profiling attacks could also be launched against anonymous users of forums and platforms. Thanks to the alleged *deanonymization attacks*, real identity is uncovered by tracking cookies, network topology, or group memberships [11].

6.3. Recognition attack

The recognition of network infrastructures is frequent in pentesting procedures and reconnaissance activities [5]. This is mostly due to IP-based connections, whose natural functioning exposes valuable details about routing algorithms, network protocols, and topologies.

Despite having network and data security protocols to guarantee confidentiality, integrity, and availability, communications exhibit inevitable traces, such as domains, IP addresses, hostnames, operating systems, or DNS records, that can be exploited [3].

These matters would not only threaten computer networks and systems, but they generate a leakage of sensitive information from individuals and companies: email addresses, private or internal network topologies, directory entries, distribution of servers, routers, IP-based devices, geolocations, active services, existing vulnerabilities, among others [3].

The attack surface alarmingly grows with the existence of smart cities, houses, and devices. They amplify the home or company connectivity exposure and manage high volumes of open data. On the other hand, AI automates scanning and

learns from vulnerability patterns, thus evading detection systems and augmenting the efficacy of intrusions [4]. Anomaly Detection methods are especially applied for these illegitimate purposes.

7. Yet, all is not lost

The culture of responsibility and awareness is conspicuous by its absence in cybersecurity. However, a person should show a responsible attitude to protect the information and avoid the amount of PII emitted to the network. In this line, we suggest for users some simple tips such as (i) minimizing virtual friendships to only known contacts, (ii) reducing the publication of personal photos, videos, and sensible information on social networks, (iii) deleting any personal details that are not strictly necessary to be publicly accessible, such as the phone number, location, or email addresses, (iv) configuring restrictive rules within privacy settings of online services and web browsers, (v) installing complementary privacy tools, such as ad-blockers and anti-tracking extensions, and (vi) being aware of data leaks that might affect personal accounts and credentials.

Privacy-compromising cyberattacks should also be tackled jointly. Some advice to mitigate collateral damage would be (i) not to publish details of other people or entities without their consent, (ii) to review the digital habits and behaviour on social networks of close contacts, and (iii) warn others of the risks beyond open data.

Nevertheless, some cyberattacks are not directly dependent on end-users. In this sense, it would be recommendable for service providers to (i) adopt privacy as a design principle at the beginning of each project, (ii) protect the data of customers and third-parties under privacy standards, (iii) review the vulnerabilities of software and the exposure of AI-based services to avoid data leaks or intrusions, (iv) make employees aware of not disclosing internal details, and (v) provide users with functionalities for the tracking, handling and cleaning of personal data.

On balance, industry, and academia should continue to develop innovative products that (i) dump only the desired information into cyberspace, (ii) track the PII of the users through the web, (iii) facilitate the removal of outdated or undesired published details, and (iv) alert end-users clearly about bad practices when engaging in online activities. These functionalities should strike a balance between protection and usability to effectively facilitate their integration into digital habits.

Finally, an effort is required to design legislative frameworks for protecting data, guaranteeing people's digital rights, and establishing strong data management regulations.

8. What's coming up

This article enumerated a significant number of threats that arise with the misuse of publicly available information, thus demonstrating that open sources might be highly compromising. We suggest a taxonomy composed of three threat dimensions and ten major types of attacks based on how open data is employed. Although we comprehensively reveal human weaknesses, we presented some recommendations to alleviate the problem.

However, it is widely assumed that the web will be even more flooded with open data, and smart technologies will increment the management of personal information. Together with the increasing autonomy and effectiveness of AI-based cyberoffence, a perfect combo to danger in modern society arises.

From such a scenario, some appealing challenges appear. One direction is to analyze and evaluate vulnerabilities in current frameworks of data management by putting the focus on AI-powered offences. We could highlight the need to detect manipulated multimedia, artificial social engineering campaigns, and coordinated phishing attacks. Another clear line of action is the research, analysis, and proposal of technical guidelines based on the *privacy by design* principle for new-generation applications to reduce the likelihood of data breach and privacy-compromising vulnerabilities. Finally, it would be worthwhile to explore those open data sources in which users unadvisedly leave personal traces and further inspect the associated privacy-compromising threats.

Acknowledgments

This study was partially funded by the Spanish Government grants FPU18/00304 and RYC-2015-18210, co-funded by the European Social

Fund.

REFERENCES

1. D. Díaz López, G. Dólera Tormo, F. Gómez Mármol, J. M. Alcaraz Calero, and G. Martínez Pérez, "Live Digital, Remember Digital: State of the Art and Research Challenges," *Computers & Electrical Engineering*, vol. 40, no. 1, pp. 109–120, 2014.
2. S. R. Sahoo and B. B. Gupta, "Classification of various attacks and their defence mechanism in online social networks: a survey," *Enterprise Information Systems*, vol. 13, no. 6, pp. 832–864, 2019.
3. J. Pastor-Galindo, P. Nespoli, F. Gómez Mármol, and G. Martínez Pérez, "The not yet exploited goldmine of osint: Opportunities, open challenges and future trends," *IEEE Access*, vol. 8, pp. 10 282–10 304, 2020.
4. M. Brundage, S. Avin, J. Clark, H. Toner, P. Eckersley, B. Garfinkel, A. Dafoe, P. Scharre, T. Zeitzoff, B. Filar, H. Anderson, H. Roff, G. C. Allen, J. Steinhardt, C. Flynn, S. Ó. hÉigeartaigh, S. Beard, H. Belfield, S. Farquhar, C. Lyle, R. Crootof, O. Evans, M. Page, J. Bryson, R. Yampolskiy, and D. Amodei, "The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation," University of Cambridge Repository, Tech. Rep. February 2018, 2018.
5. N. Kaloudi and J. Li, "The AI-based cyber threat landscape: A survey," *ACM Computing Surveys*, vol. 53, no. 1, 2020.
6. SophosLabs, "Sophos 2020 Threat Report. We're covering your blind spots." Sophos, Tech. Rep., 2019.
7. Check Point Research, "Cyber Security Report 2020," Check Point, Tech. Rep., 2020.
8. M. M. Yamin, M. Ullah, H. Ullah, and B. Katt, "Weaponized ai for cyber attacks," *Journal of Information Security and Applications*, vol. 57, p. 102722, 2021.
9. J. Pastor-Galindo, M. Zago, P. Nespoli, S. L. Bernal, A. H. Celdrán, M. G. Pérez, J. A. Ruipérez-Valiente, G. M. Pérez, and F. G. Mármol, "Spotting political social bots in twitter: A use case of the 2019 spanish general election," *IEEE Transactions on Network and Service Management*, vol. 17, no. 4, pp. 2156–2170, 2020.
10. R. Derbyshire, B. Green, D. Prince, A. Mauthe, and D. Hutchison, "An analysis of cyber security attack taxonomies," in *2018 IEEE European Symposium on Security and Privacy Workshops*, 2018, pp. 153–161.
11. M. Fire, R. Goldschmidt, and Y. Elovici, "Online social networks: Threats and solutions," *IEEE Communications Surveys and Tutorials*, vol. 16, no. 4, pp. 2019–2036, 2014.
12. T.King, N.Aggarwal, M. Taddeo, and L.Floridi, *Artificial Intelligence Crime: An Interdisciplinary Analysis of Foreseeable Threats and Solutions*. Springer Netherlands, 2020, vol. 26.
13. M. Edwards, R. Larson, B. Green, A. Rashid, and A. Baron, "Panning for gold: Automatically analysing online social engineering attack surfaces," *Computers and Security*, vol. 69, pp. 18–34, 2017.
14. K. Krombholz, H. Hobel, M. Huber, and E. Weippl, "Advanced social engineering attacks," *Journal of Information Security and Applications*, vol. 22, pp. 113–122, 2015.
15. F. Salahdine and N. Kaabouch, "Social engineering attacks: A survey," *Future Internet*, vol. 11, no. 4, 2019.

Javier Pastor-Galindo received B.Sc. and M.Sc. degrees in Computer Science from the University of Murcia, Spain. He is currently a PhD student in the same institution with an FPU contract granted by the Spanish Ministry of Universities. His research interests focus on open source intelligence (OSINT), security, and privacy.
Contact: javierpg@um.es

Félix Gómez Mármol received a M.Sc. and Ph.D. in Computer Science from the University of Murcia. He is currently a researcher in the Department of Information and Communications Engineering at the University of Murcia, Spain. His research interests include cybersecurity, internet of things, machine learning and bio-inspired algorithms.
Contact: felixgm@um.es

Gregorio Martínez Pérez received a Ph.D. degree in Computer Science at the University of Murcia, where he is Full Professor since 2014. His scientific activity is mainly devoted to cybersecurity and data science. He is working on different national and European IST research projects related to these topics, being Principal Investigator for UMU in most of them.
Contact: gregorio@um.es