A Repository of Actionable Information on the Internet of Things

Anna Felkner[a] and Marcin Rytel[b]
NASK - Research and Academic Computer Network, Kolska 12, 01-045 Warsaw, Poland

Keywords: Internet of Things, IoT, Security, Vulnerability, Vulnerability Database, Exploits.

Abstract: This paper describes efforts to improve the security of the Internet of Things world. One of the many problems faced by the users, but also producers or owners of networks or people who deal with cybersecurity in a broad sense on a daily basis, such as employees of CSIRTs, is the issue of vulnerabilities in devices or software. Even though the most serious vulnerabilities are more commonly presented in the news, still the vast majority of them is only known to cybersecurity specialists and not to the users who own the vulnerable devices and therefore can be in danger. The results of our recent research show that there is still no satisfactory source of information about vulnerabilities in Internet of Things devices and software. This is why we decided to create such a repository (variotdbs.pl) where information about vulnerabilities and exploits targeting IoT devices can be easily found by everyone.

1 INTRODUCTION

Given that Internet of Things (IoT) devices accompany us every day, both in private and professional life, both at home and in industry, in healthcare and on the streets, we cannot ignore the issues related to their security. Fortunately, in recent times, the awareness of users of these devices has been increasing, but it is still not a completely satisfied topic.

In our work, we focused on looking for information about exploits and vulnerabilities in the Internet of Things and noticed that there is no single source that would present a wide range of information related to this aspect of security. From our research, we found that while information is available online, there has not been a single service offering IoT data to date. National vulnerability databases contain some IoT entries, but lack mechanisms to distinguish them from other vulnerabilities. Moreover, information about a lot of vulnerabilities pertaining to the Internet of Things world never makes it into these databases, but can be found scattered across the Internet. We therefore decided to create such a source.

To start with, we analysed more than one hundred unique sources of different types. Structured sources which contain information not only about IoT but also, or rather mainly about general IT, such as vendor bulletins, various national vulnerability databases. Apart from those, also unstructured sources such as reports, blogs or individual websites were analysed.

In the paper (Rytel et al., 2020), we presented an overview summarizing our effort to identify and evaluate publicly available sources of vulnerability information, focusing on their usefulness in the IoT domain. After analysing the sources, the next step towards providing not only users, but also CSIRTs (Computer Security Incident Response Team) or network owners, with actionable information on the Internet of Things was to create a database with this information. In the paper (Janiszewski et al., 2021), we presented the results of our research aimed at building such a database, i.e., how to obtain, standardize, aggregate and correlate vulnerability information, as well as how to enrich and select IoT-related data. We have derived and demonstrated that existing databases provide very different ranges of information and for this reason, there is no single comprehensive source of information. Moreover, different sources present vulnerability information at different times - some sooner others later, and the differences in publication dates of the information are significant. Our results show that aggregating information from different sources can be very beneficial and can potentially increase the value of information for taking action. We have also shown that introducing some more advanced concepts like trust management and AI-based meta-information extraction, can provide a higher level of information completeness, as well as

[a] https://orcid.org/0000-0003-3813-4840
[b] https://orcid.org/0000-0002-8590-8565

Felkner, A. and Rytel, M.
A Repository of Actionable Information on the Internet of Things.
DOI: 10.5220/0011307100003286
In *Proceedings of the 19th International Conference on Wireless Networks and Mobile Systems (WINSYS 2022)*, pages 69-75
ISBN: 978-989-758-592-0; ISSN: 2184-948X

assess the usefulness and reliability of the data.

The rest of the paper is structured as follows. The second chapter provides an overview of related work. The third chapter presents the process of creating the database. The fourth chapter focuses on presenting the results of our work and the fifth chapter describes the variotdbs.pl (NASK, 2022) website, where the results of our work are published and which anyone can use for their own applications. The sixth chapter focuses on the summary and presents our ideas for further work.

2 RELATED WORKS

IoT security is a major focus of research programs and plans at many levels, including the "Input to the Horizon Europe Programme 2021-2027 Priorities for the definition of a Strategic Research and Innovation Agenda in Cybersecurity" (ECSO WG 6 – SRIA and Cybersecurity Technologies, 2020). National strategies are even more focused on the topic and explicit on this issue. In the document (Felkner et al., 2021), which presents an analysis on the European and Japanese side, it is shown that the Internet of Things is one of the most frequently addressed issues in European national cybersecurity strategies and is also one of the areas of common interest between the European Union and Japan, indicating that this issue is globally relevant.

Having information about IoT devices vulnerabilities is critical from the perspective of device owners, service providers, network owners, and device producers. Obtaining this information is also critical from national and sector CSIRTs perspective. Vulnerability management is one of the key aspects of security in both the IT and IoT or IIoT (industrial IoT) worlds. Risk assessment at different levels can also be done on the basis of vulnerability management, as shown for example in the article (Janiszewski et al., 2019). The survey of IoT vulnerability data sources was described in (Rytel et al., 2020), while in (Janiszewski et al., 2021) you can find out how the process of creating database was performed.

Building such a database is definitely a non-trivial task. This can be seen from the few attempts that have been made to build such databases. One example was a small-scale attempt to create such a database, which was undertaken at the University of Central Florida, but unfortunately this database is not currently available. Its description can be found here (Ling et al., 2017), but its dataset was not publicly available. A second promising approach was the design of a database of vulnerabilities and attacks on IoT infrastructure, presented by researchers at the University of New South Wales in Australia (Nerwich et al., 2020). According to the creators, this database supported integration with other vulnerability databases such as the National Vulnerability Database (NVD) and provided an API to access the data for integration with other applications. Its goal was to serve as a knowledge base for IoT application developers and security researchers. Unfortunately, the database has not been made public anywhere and the only information about it can be found in the above-mentioned article, which allows us to conclude that the work on the database is not continued, but was only a Proof of Concept. This still leaves a gap that those responsible for the security of IoT devices should fill. Therefore, this article presents the whole process of creating and publishing such a database.

This paper is written based on the results obtained during the work performed in the Vulnerability and Attack Repository for IoT project (VARIoT, 2022). The objective of the project is to provide actionable information about Internet of Things devices that can be processed manually or automatically to ensure the cybersecurity of these devices. This involves not only creating and presenting a database of information about exploits and vulnerabilities in the Internet of Things, but also, among other things, scanning the Internet to identify vulnerable, publicly available IoT devices. Laboratories have also been built to test both legitimate and malicious IoT traffic, IoT artifacts, and IoT anomaly models. Aggregated and anonymous statistics on infected and vulnerable IoT devices will also be prepared in the near future. All these tasks are carried out in collaboration with our partners, namely Stichting The Shadowserver Foundation Europe, Security Made In Letzebuerg G.I.E., Institut Mines-Télécom and Mondragon Goi Eskola Politeknikoa Jose Maria Arizmendiarrieta S COOP.

3 DATABASE CREATION

In our work, we only analyse publicly available, free sources of information, which excludes paid services such as Vulners vulnerability and exploit aggregator (Vulners, nd), among others. In addition to the entries collected from structured sources listed in Table 1, we also look for recent posts and articles found over the Internet. For these, the relevant metadata can be extracted from the raw text. One of the unique features of the built VARIoT vulnerability database is the correlation and aggregation of vulnerability information from various publicly available sources.

As mentioned earlier, there are many publicly

Table 1: Used information sources.

Short Name	Full Name	Type of db	Reference
BID	SecutiryFocus Bugtraq	Vuln/Expl	(BID, nd)
CERT CC	Carnegie Mellon Univ. CERT Coordination Center	Vuln	(CERT, nd)
CNNVD	Chinese National Database of Information Security	Vuln	(CNNVD, nd)
CNVD	China National Vulnerability Database	Vuln	(CNVD, nd)
ExploitDatabase	EXPLOIT-DATABASE.NET	Expl	(ExploitDb, nd)
Exploit-DB	Exploit Database by Offensive Security	Expl	(Exploit-DB, nd)
ICS-CERT CN	Chinese ICS-CERT website	Vuln	(ICS-CERT, nd)
IVD	ICS Vulnerability Database	Vuln	(IVD, nd)
JVNDB	Japan Vulnerabilities Notes Database	Vuln	(JVNDB, nd)
NVD	National Vulnerability Database	Vuln	(NVD, nd)
Packet Storm	Packet Storm Security	Vuln/Expl	(PacketStorm, nd)
Vulmon	Vulmon Vulnerability Search Engine	Vuln	(Vulmon, nd)
VUL-HUB	VUL-HUB Information Security Vuln. Portal	Vuln	(VUL-HUB, nd)
ZDI	Zero Day Initiative	Vuln	(ZDI, nd)
ZSL	Zero Science Lab	Vuln	(ZSL, nd)

available databases containing various information about vulnerabilities in different types of hardware and software. Just a few of them are dedicated solely to Internet of Things or at least point to such vulnerabilities in some way, but none of them directly aggregate information from other sources. The vulnerability database creation is show in Figure 1.

The entire process of creating the database of IoT vulnerabilities and exploits that we prepared within the VARIoT project is shown in Figure 2 and can be briefly described as follows. In the first step, identification and selection of valuable sources of information related to vulnerabilities and exploits is performed. Many types of sources are of interest, both structured, such as national vulnerability databases (from which it is easier to retrieve data) and unstructured, such as blogs, individual websites (from which retrieving data is usually more complicated but may provide information ahead of official databases or contain completely unique data). In the second step, we harvest information from the sources and store it in so-called raw databases. In the third step, the information is standardized - e.g., the names of relevant fields are unified and complementary information is added. In this way so-called low databases are created. In the fourth step, information from various sources about a given vulnerability or exploit is correlated and aggregated. Based on the results of this process, a medium database is created that contains all the information from every low database, and each entry in this database corresponds to one vulnerability or exploit. Each field in the entry contains information from the corresponding fields in the low databases. The fifth step is to enhance and select the most reliable information about each vulnerability and exploit. This process uses two mechanisms: metainformation extraction and trust management. In order to enable the extraction of metainformation in an automatic way, we applied various mechanisms of artificial intelligence and natural language processing. Based on the information contained in the database, we prepared dictionaries of information about producers, models and device types as well as about vulnerability types. These dictionaries were used as keywords for searching in a text and as training data sets for other methods. The trust evaluation aims to select the most reliable and informative part of the information, evaluate the reliability of the information, and identify IoT-related vulnerabilities and exploits. It is done based on source reputation, convergence of information from different sources, aggregation and classification method, and additional searches. The result of this process is the creation of a high database. The database created as described above can then be shared and used by different entities for different purposes. IoT-related information is also carried out by filtering at different levels using IoT device taxonomy we have created, an internal IoT device catalogue, a keyword-based filtering mechanism, etc. Each step is described in detail in the paper (Janiszewski et al., 2021).

4 RESULTS

As of this writing, there are 1 151 723 entries in the low databases, resulting in 207 673 entries in the medium and high database, of which 24 969 are IoT-related (as of April 10, 2022). About 80% of all entries that are in the high database contain data re-

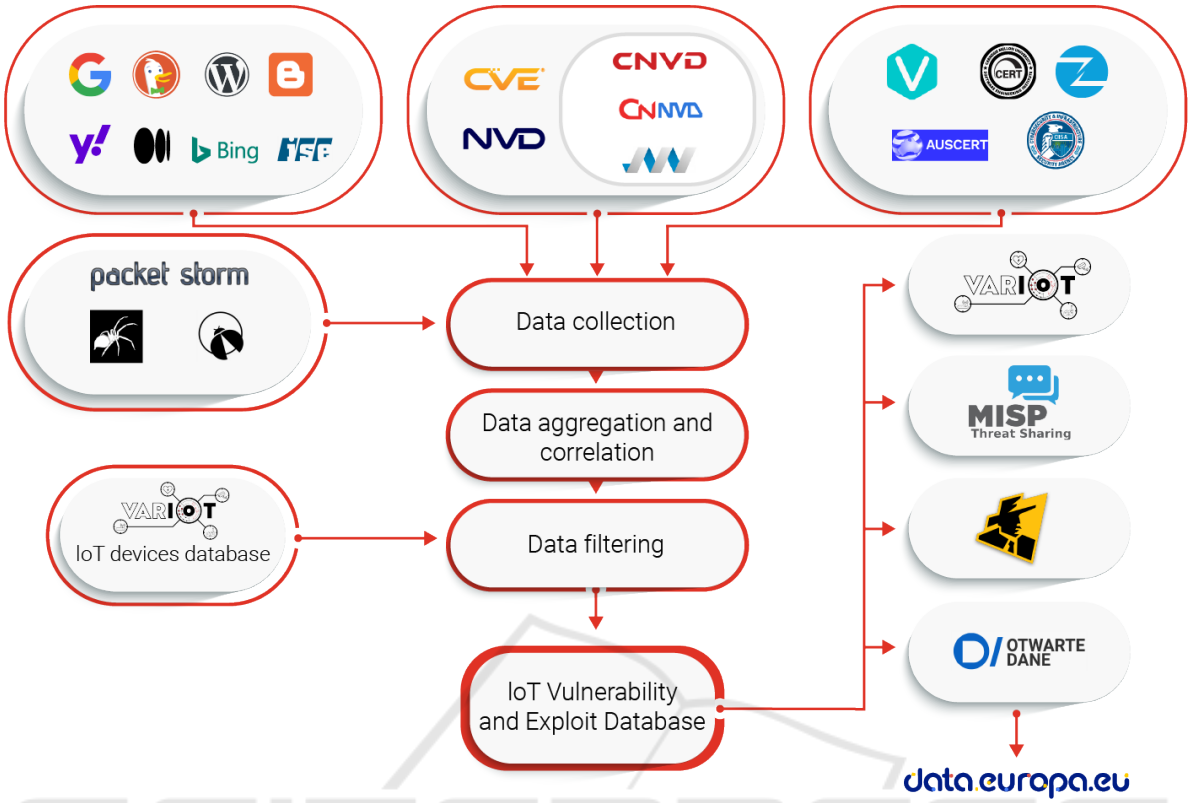

Figure 1: Vulnerability database creation.

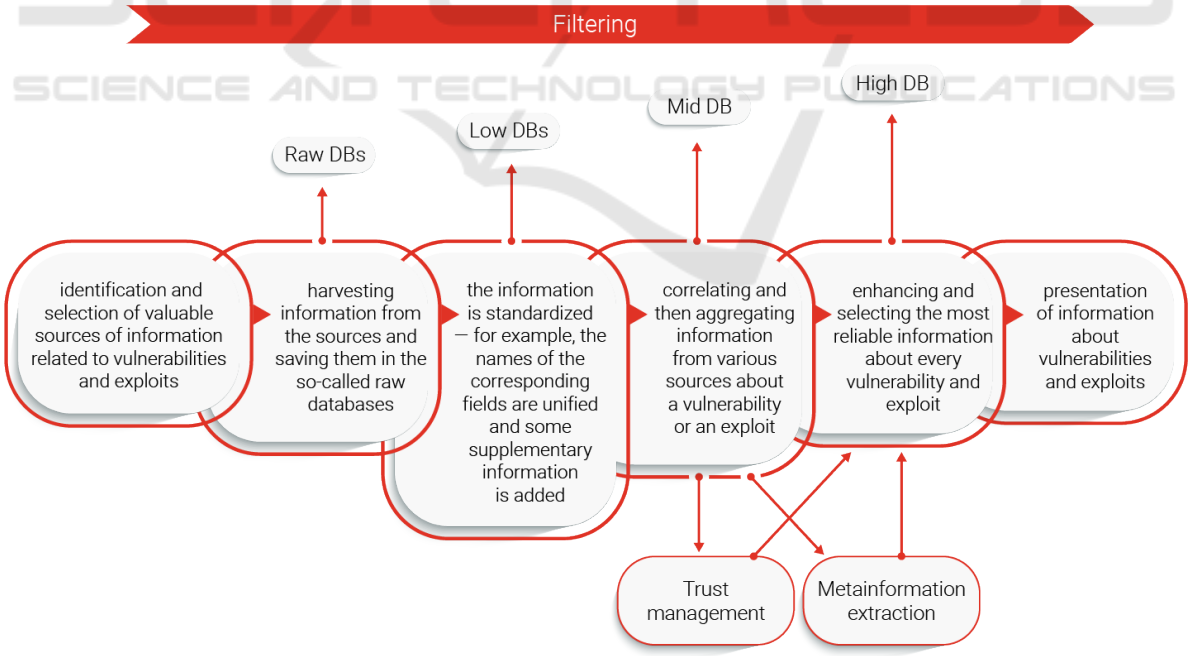

Figure 2: The process of repository creation.

trieved from more than one information source.

Creating a structured database, publicly available, with information about known technical vulnerabilities and exploits is extremely beneficial to all stakeholders: users, producers and network owners, as well as CSIRTs or other people who work on device and

software security. We have studied a large cross-section of different types of sources, and from this we can conclude that by collecting data from a variety of sources, we can obtain a more complete and comprehensive entry on a given vulnerability or exploit than from a single source. Because we search many different types of publicly available sources and our database correlates and aggregates data from these sources, this makes each entry rich in information that can be used to ensure the security of IoT devices, and reduces the risk of missing some data or delays in obtaining information about specific vulnerabilities.

The number of entries in each database can be seen in Table 2.

Table 2: Number of entries in databases.

Database	Number of entries
BID	100 825
CERT CC	3 555
CNNVD	183 081
CNVD	174 881
ExploitDatabase	98 877
Exploit-DB	44 854
ICS-CERT CN	2 929
IVD	3 449
JVNDB	139 536
NVD	183 022
Packet Storm	154 228
Vulmon	40 305
VUL-HUB	11 466
ZDI	9 905
ZSL	810

As this is a short form paper we only focus on the most important results. More information and statistics related to the results can be found in the paper (Janiszewski et al., 2021). The document is from a year ago but it shows more or less what the relationships between the databases are, statistics, ranges and more.

5 PUBLICATION OF A DATABASE OF INFORMATION ON VULNERABILITIES AND EXPLOITS

One of the main results of our work is a database containing information about vulnerabilities and exploits in the Internet of Things found on the Internet. This database is published at: https://www.variotdbs.pl/.

The main data presented is organized into three sections:

- Vulnerabilities - here security vulnerabilities affecting IoT devices are shown.
- Exploits - publicly available exploits targeting IoT devices are shown here.
- News - here you can see an automatically generated news feed about IoT security crawled using our custom search engine.

There are also two additional sections on the website:

- API - here is a description of how to easily retrieve the data we provide on the website through the API. This can be done using one of two well structured file formats: JSON and JSON-LD.
- Ontology - here is a description of the ontology of the VARIoT vulnerability database entries.

Vulnerability and exploit databases are built on the basis of the previously described sources. The news feed uses many available search engines and creates an entry based on the information found in this way. Figure 3 shows a diagram of how the search engine works. Vulnerability and news feed sections use solutions based on natural language processing (NLP), machine learning (ML) and artificial intelligence (AI) to better tailor the entry and combine data obtained from multiple sources.

The developed mechanism is able to extract information from unstructured news sources such as blogs, reports and articles. Moreover, it calculates the trust to the selected information to better evaluate its relevance. In the context of vulnerabilities and exploits, trust is based on source reliability scores established based on our knowledge about these sources, and in the context of news feeds, trust is based on information extracted from found entries, i.e. keywords, manufacturers and product names, vulnerability types, and links to well-known vulnerabilities databases.

The *Vulnerabilities* section allows you to view the latest IoT vulnerabilities and search for them. Vulnerability searches are possible using a variety of attributes, whether by vendor, device model, version, CVE (Common Vulnerabilities and Exposures), vulnerability identifier, CWE (Common Weakness Enumeration), type, or by a phrase describing the vulnerability. Such a search gives many possibilities, as it is possible to find a description of vulnerabilities using any of the above fields, e.g. device type, and other related sources of information about vulnerabilities. Each entry consists of data on a particular vulnerability found from various sources. Each entry includes sources of information and calculated confidence levels, as well as aggregated links to external sources that can be consulted for further information.

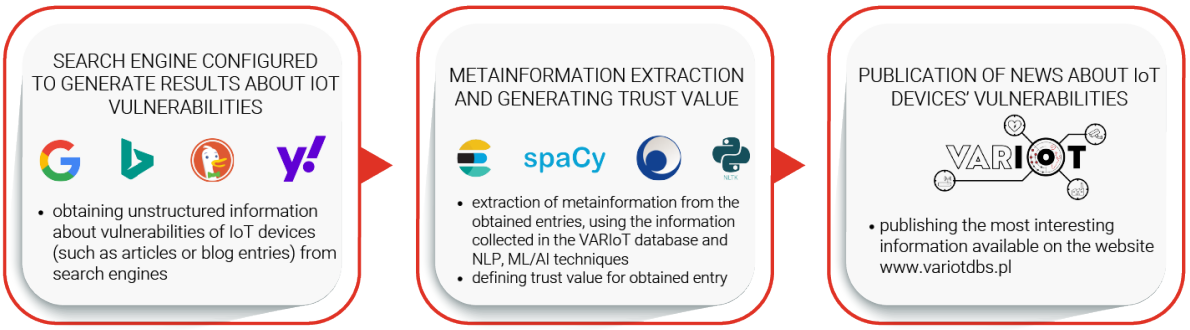

Figure 3: IoT vulnerability search engine.

The *Exploit* section allows you to view exploits that may pose a threat to IoT devices. Currently (as of April 5, 2022), it is in the testing phase, but by the time this article is published, it will have been refined and will work in a similar way to the vulnerability section.

The *News* section shows different types of information related to vulnerabilities in the world of the Internet of Things. The news are collected from various, mostly unstructured sources such as reports, blogs, information provided by people who work with vulnerability research or in any way related to cybersecurity. The news is retrieved using our scripts to filter the search results. Additional information about vulnerabilities, attacked products and external identifiers is obtained using NLP and custom filters from found news.

6 CONCLUSIONS AND FUTURE WORKS

One of the most important conclusions from our research is that our vulnerabilities and exploits database fits well in the gap of the IoT devices security.

The mechanisms implemented in our work and the information they provided can form the basis for building different types of services. The database we created can be used by vulnerability scanners as a repository of information about IoT vulnerabilities. This is because the information we collect is in various aspects richer, broader and more comprehensive than that currently used by popular vulnerability scanners, especially in the IoT context.

In the future, we are also considering another very practical service, namely providing for a given product name (specific producer and model) a list of possible vulnerabilities. This extension seems to be easy to build and at the same time extremely useful. The user will be able to enter the name of their device and will get a list of all known vulnerabilities that affect it. It seems to be even more useful from the producer's, network owner's or CSIRT operator's point of view. Previously, an inventory of IoT assets would have been necessary, but this approach can yield much better results than scanning with vulnerability scanners, which will not necessarily be able to identify the device well. We see great potential for our solution.

The repository that we prepared will be publicly accessible through the European Data Portal - data.europa.eu and through national Data Portals (such as the Polish Open Data Portal (OpenData, 2022)), as well as other sources, such as the Malware Information Sharing Platform (MISP), which is widely used by the cybersecurity analyst community, or through Shadowserver's free daily remediation feeds.

ACKNOWLEDGEMENTS

Scientific work published as part of an international project co-financed by the Connecting Europe Facility of the European Union, TENtec n. 28263632 and by the program of the Minister of Science and Higher Education entitled "PMW" in the years 2020–2022; contract No. 5095/CEF/2020/2.

REFERENCES

BID (n.d.). Secutiryfocus Bugtraq database. https://www.securityfocus.com/bid Last accessed on 2021-04-12.

CERT (n.d.). Carnegie Mellon University CERT Coordination Center. https:// www.kb.cert.org/vuls/ Last accessed on 2022-04-12.

CNNVD (n.d.). Chinese National Vulnerability Database of Information Security. http://www.cnnvd.org.cn/ Last accessed on 2022-04-12.

CNVD (n.d.). China National Vulnerability Database. https://www.cnvd.org.cn/ Last accessed on 2022-04-12.

ECSO WG 6 – SRIA and Cybersecurity Technologies (2020). Input to the horizon europe programme 2021-2027. priorities for the definition of a strategic research and innovation agenda in cybersecurity. https://ecs-org.eu/documents/publications/5fdc4c5deb6f9.pdf. Last accessed on 2022-04-12.

Exploit-DB (n.d.). Offensive Security's Exploit Database Archive. https://www.exploit-db.com/ Last accessed on 2022-04-12.

ExploitDb (n.d.). EXPLOIT-DATABASE.NET. https://www.exploit-database.net/ Last accessed on 2022-04-12.

Felkner, A., Kadobayashi, Y., Janiszewski, M., Fantin, S., Ruiz, J. F., Kozakiewicz, A., and Blanc, G. (2021). *Cybersecurity Research Analysis Report for Europe and Japan: Cybersecurity and Privacy Dialogue Between Europe and Japan*, volume 75 of *Studies in Big Data*. Springer International Publishing, Cham.

ICS-CERT (n.d.). Chinese ICS-CERT website. https://www.ics-cert.org.cn/portal/index.html Last accessed on 2022-04-12.

IVD (n.d.). Ics Vulnerability Database. http://ivd.winicssec.com/ Last accessed on 2022-04-12.

Janiszewski, M., Felkner, A., and Lewandowski, P. (2019). A Novel Approach to National-level Cyber Risk Assessment Based on Vulnerability Management and Threat Intelligence. *Journal of Telecommunications and Information Technology*, 2:5–14.

Janiszewski, M., Felkner, A., Lewandowski, P., Rytel, M., and Romanowski, H. (2021). Automatic actionable information processing and trust management towards safer internet of things. *Sensors*, 21(13):4359.

JVNDB (n.d.). Japan Vulnerabilities Notes Database. https://jvndb.jvn.jp/en/ Last accessed on 2022-04-12.

Ling, Z., Liu, K., Xu, Y., Jin, Y., and Fu, X. (2017). An end-to-end view of iot security and privacy. In *GLOBECOM 2017 - 2017 IEEE Global Communications Conference*, pages 1–7.

NASK (2022). https://www.variotdbs.pl/, Last accessed on 2022-04-12.

Nerwich, M., Gauravaram, P., Paik, H.-y., and Nepal, S. (2020). Vulnerability database as a service for iot. In Batina, L. and Li, G., editors, *Applications and Techniques in Information Security*, pages 95–107, Singapore. Springer Singapore.

NVD (n.d.). National Vulnerability Database. https://nvd.nist.gov/vuln/search Last accessed on 2022-04-12.

OpenData (2022). Polish Open Data Portal. https://dane.gov.pl/en, Last accessed on 2022-04-12.

PacketStorm (n.d.). Packet Storm. https://packetstormsecurity.com/ Last accessed on 2022-04-12.

Rytel, M., Felkner, A., and Janiszewski, M. (2020). Towards a Safer Internet of Things—A Survey of IoT Vulnerability Data Sources. *Sensors*, 20(21):5969.

VARIoT (2019-2022). Vulnerability and Attack Repository for IoT project. https://www.variot.eu Last accessed on 2022-04-12.

VUL-HUB (n.d.). Vul-hub Information Security Vulnerability Portal. http://www.cve.scap.org.cn/ Last accessed on 2022-04-12.

Vulmon (n.d.). Vulmon Vulnerability Search Engine. https://vulmon.com/ Last accessed on 2022-04-12.

Vulners (n.d.). Vulners - Vulnerability Data Base. https://vulners.com/ Last accessed on 2022-04-12.

ZDI (n.d.). Zero Day Initiative. https://www.zeroday initiative.com/, Last accessed on 2022-04-12.

ZSL (n.d.). Zero Science Lab. https://www.zeroscience.mk/en/index.php, Last accessed on 2022-04-12.