Search Query:
("threat intelligence" OR "cyber threat intelligence") AND ("underground forums" OR "darknet" OR "Twitter" OR "Telegram" OR "Discord" OR "OSINT" OR "Open source intelligence" OR "social media" OR "security reports")
Papers starting from 2017

No	Google Scholar		Number of Hits				100	
1	Main Author	Year	Title	Brief summary	Comment/Link	Arxiv Paper	Date of Search	Excluded?
3	H. Shin	2021	# twiti: Social listening for threat intelligence	Automated NLP model to extract early and unique malware IOC's from twitter			19.02.2024	
5	Victor Adewopo	2020	Exploring Open Source Information for Cyber Threat Intelligence	System with 82% accuracy for CTI extraction. On Twitter and HackerForums and analysis of 8'000 reported cases in US and do risk profiling for geo-location cyberattacks			19.02.2024	
6	J. Zhao	2020	TIMiner: Automatically extracting and analyzing categorized cyber threat intelligence from social data	CTI extraction from social media data. focus on uncseen IoC and domain tags (finance or govt.)			19.02.2024	
7	P. Koloveas	2021	inTIME: A Machine Learning-Based Framework for Gathering and Leveraging Web Data to Cyber-Threat Intelligence	ML-approach to identify, collect, analyse, extract, integrate, and share CTI from different sources (clear/dark web)			19.02.2024	
9	Uğur Tekin	2021	Obtaining Cyber Threat Intelligence Data From Twitter With Deep Learning Methods	DL methods to extract CTI from twitter and classified			19.02.2024	
14	Linn-Mari Kristiansen	2020	CTI-Twitter: Gathering Cyber Threat Intelligence from Twitter using Integrated Supervised and Unsupervised Learning	collect, process, analyze CTI from twitter			19.02.2024	
15	Ariel Rodriguez	2020	Enhancing data quality in real-time threat intelligence systems using machine learning	Methods to avoid keyword filtering of CTI extraction. Based on clustering and removing of unimportant clusters			19.02.2024	
20	Rui Azevedo	2019	PURE: Generating Quality Threat Intelligence by Clustering and Correlating OSINT	IoCs are correlated together, this method is built on top of already existing OSINT feeds			19.02.2024	
21	Md Imran Hossen	2021	Generating cyber threat intelligence to discover potential security threats using classification and topic modeling	They collect data, use different methods to classify	Chapter of a book		19.02.2024	
24	Yongyan Guo	2023	A framework for threat intelligence extraction and fusion	Using the idea of a Cybersecurity Knowledge Graph (CKG) the authors introduce a framework which claim to better extract CTI with OSINT. There is a focus on Entity extraction and entity relation, which are usually divided subtasks but are here fused together.			23.02.2024	
26	Avishek Bose	2021	Tracing Relevant Twitter Accounts Active in Cyber Threat Intelligence Domain by Exploiting Content and Structure of Twitter Network	Automated method to extract twitter users that post relevant CTI info			23.02.2024	
32	Masashi Kadoguchi	2020	Deep Self-Supervised Clustering of the Dark Web for Cyber Threat Intelligence	ML for CTI extraction from specific darkweb posts (no clustering mentioned in the abstract			23.02.2024	
33	Yinghai Zhou	2022	CTI View: APT Threat Intelligence Analysis System	Development of CTI View focus on: 1) Heterogeneous inputs 2) IOC/TTP extraction 3) Entity extraction (labeling?)			23.02.2024	
34	Georgios Sakellariou	2023	SECDFAN: A Cyber Threat Intelligence System for Discussion Forums Utilization	SECDFAN reference architecture. A framework for CTI product sharing			23.02.2024	
37	Nolan Arnold	2019	Dark-Net Ecosystem Cyber-Threat Intelligence (CTI) Tool	A CTI tool uses a social network to find cyber-threats on major Dark-net sources. It helps quickly spot emerging threats so that security measures can be taken promptly, either proactively or reactively.			23.02.2024	
42	Andrea Tundis	2022	A Feature-driven Method for Automating the Assessment of OSINT Cyber Threat Sources	method to automate the assessment of cyber threat intelligence sources and predict a relevance score for each source is proposed.			23.02.2024	
45	Jeonghun Cha	2020	Blockchain-Based Cyber Threat Intelligence System Architecture for Sustainable Computing	CTI sharing platform which is efficent in terms of memory reliability, privacy, scalability, and sustainability.			23.02.2024	
46	Benjamin Ampel	2020	Labeling Hacker Exploits for Proactive Cyber Threat Intelligence: A Deep Transfer Learning Approach	framework for the automated collection and categorization of hacker forum exploit source code.	Link		24.02.2024	

Procedure:
1. We went throught the first 10 pages hits on Google Scholar for the query (on the left).
2. We reviewed the title and the abstract to examine whether the paper is of interest to us.
3. We identified 71 paper, and excluded 21 papers that are older than 5 years - published prior to 2019), 2 papers because they are only arxiv preprints, and 21 papers after a detailed review (not in scope of our review).

Note: Our focus lies on the CTI extraction from a diverse range of data source (i.e. generate CTI based on the available date) - we do not focus on CTI sharing or on the detection of the identified threats in organizations.

47	Moumita Das Purba	2023	Extracting Actionable Cyber Threat Intelligence from Twitter Stream	approach to extract technical manifestations of attacks from tweets. Focus is on bettering context surrounding IoC	Link	24.02.2024
48	Tianfang Sun	2021	An Automatic Generation Approach of the Cyber Threat Intelligence Records Based on Multi-Source Information Fusion	automatic approach to generate the CTI records based on multi-type OSTIPs (GCO), combing the NLP method, machine learning method, and cybersecurity threat intelligence knowledge.	Link	24.02.2024
49	Dr. Fahim Sufi	2023	A New Social Media-Driven Cyber Threat Intelligence	CTI extraction from twitter with NLP applied on multiple different languages	Link	24.02.2024
53	Shin-Ying Huang	2020	Monitoring Social Media for Vulnerability-Threat Prediction and Topic Analysis	Development of a dynamic vulnerability-threat assessment model to predict the tendency to be exploited for vulnerability entries listed in Common Vulnerability Exposures, and also to analyze social media contents such as Twitter to extract meaningful information.	Link	24.02.2024
55	A Zenebe	2019	Cyber Threat Discovery from Dark Web	descriptive analytics and predicative analytics using machine learning on forum posts dataset from darknet to discover valuable cyber threat intelligence. Also build classifiers for exploit types and atterns and trends on hackers' plan	Link	24.02.2024
57	Fahim Sufi	2023	A global cyber-threat intelligence system with artificial intelligence and convolutional neural network	system provides critical analytical capability in the cyber-threat spectrum and uses sophisticated AI-based algorithms for anomaly detection, prediction, sentiment analysis, location detection, translation, etc.	Link	24.02.2024
62	Keisuke Furumoto	2021	Extracting Threat Intelligence Related IoT Botnet From Latest Dark Web Data Collection	Crawler adapt for the darkweb	Link	24.02.2024
68	Apurv Singh Gautam	2019	Hacker Forum Exploit and Classification for Proactive Cyber Threat Intelligence	AI approach using neural networks to automatically classify hacker forum data into predefined categories and develop interactive visualizations	Link	24.02.2024
70	Panos Panagiotou	2021	Towards Selecting Informative Content for Cyber Threat Intelligence	supervised machine learning-based text classification techniques, trained on a new dataset created for the purposes of selecting only pages with relevant CTI	Link	24.02.2024