Mobile Forensics, Big Data and Artificial Intelligence: Current Status, Challenges and Future Directions

ALEXANDROS VASILARAS
Police Major
Head of Digital Forensics Department
Hellenic Police Forensic Science Division

ILIAS PANAGIOTOPOULOS
Police Captain
Post-Doctoral Researcher
Hellenic Police Forensic Science Division

NIKOLAOS PAPADOUDIS
Police Lieutenant
Digital Evidence Examiner
Hellenic Police Forensic Science Division

CEPOL 79/2022 European Research and Science Conference
8-10 June 2022, MRU, Vilnius, Lithuania

ACKNOWLEDGEMENTS

- Alexandros Vasilaras

Police Major, Head of Digital Forensics Department, Hellenic Police Forensic Science Division. PhD candidate, Department of Telematics and Informatics, Harokopio University of Athens.

- Ilias Panagiotopoulos

Police Captain, Hellenic Police Forensic Science Division. Post-doctoral Researcher, Department of Telematics and Informatics, Harokopio University of Athens.

- Nikolaos Papadoudis

Police Lieutenant, Digital Evidence Examiner, Hellenic Police Forensic Science Division. Postgraduate Student, Department of Computer Science and Engineering, European University Cyprus.

HELLENIC POLICE FORENSIC SCIENCE DIVISION
DIGITAL FORENSICS DEPARTMENT

Hellenic Police Forensic Science Division (F.S.D.) is the National Forensic Service of Greece and provides significant scientific support and assistance to the work of the Police, but also to the work of all Prosecuting Authorities and Law Enforcement Agencies.

CEPOL 79/2022 European Research and Science Conference
8-10 June 2022, MRU, Vilnius, Lithuania

HELLENIC POLICE FORENSIC SCIENCE DIVISION
DIGITAL FORENSICS DEPARTMENT

- Specialized personnel (30 certified Digital Forensic Experts)
- What we do:
  - data recovery
  - decryption
  - examination
  - analysis
  - correlations
  - reporting

Mobile Forensics, Big Data and Artificial Intelligence: Current Status, Challenges and Future Directions

Structure of the Paper

1. Introduction
2. Mobile Investigations and Digital Forensics
3. Issues in Mobile Forensics
4. Mobile Forensics and AI Solutions
5. Conclusions

CEPOL 79/2022 European Research and Science Conference
8-10 June 2022, MRU, Vilnius, Lithuania

INTRODUCTION

The present study aims to explore:

- **the status of Mobile Forensics** in relation with **Big Data and AI**
- **the current challenges** in Mobile Forensics
- **the corresponding applications and solutions** AI provides to the challenges that investigators face
- **the legal issues and technology integration** of AI functions and software
- **the particularly useful AI research topics** that would benefit Mobile Forensics in the long term.

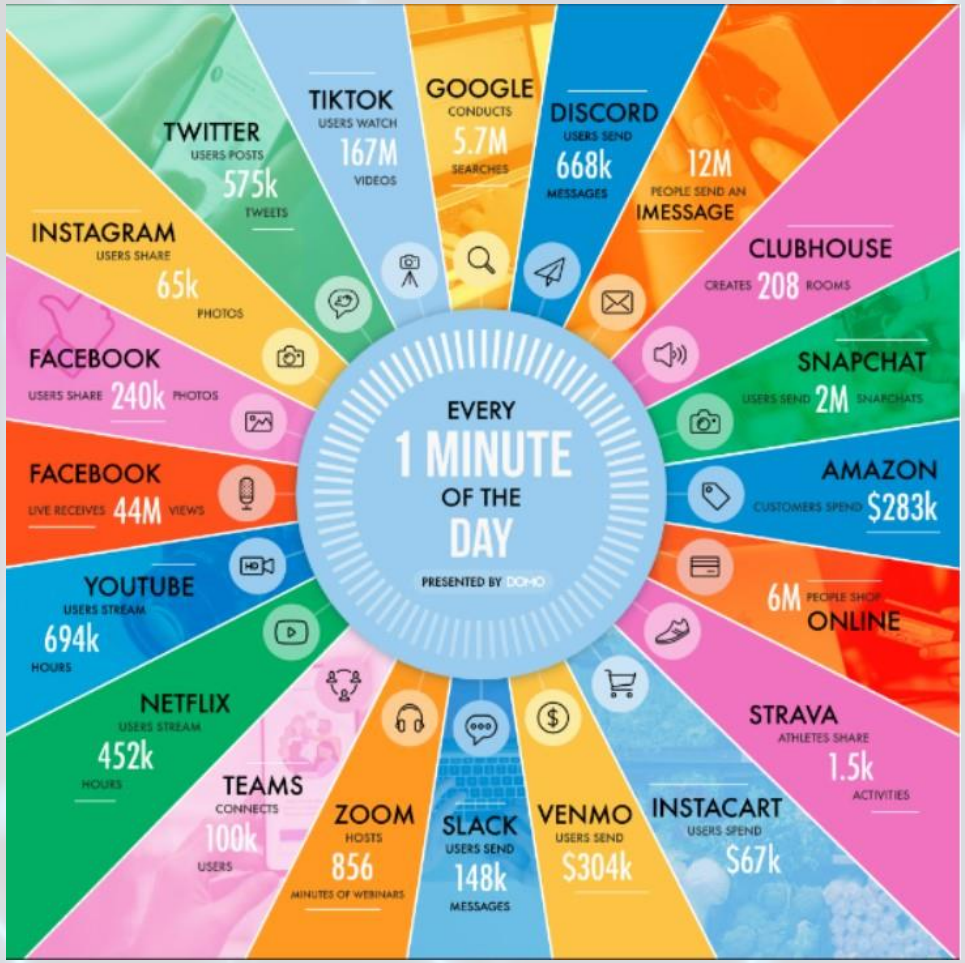
MOBILE INVESTIGATIONS AND DIGITAL FORENSICS

Digital forensics has grown rapidly due in part to the **increase in mobile devices**.

Forensic investigators face **numerous challenges dealing with digital evidence obtained from mobile devices**, which are correlated with:

- **Cloud Computing**
- **Internet of Things**
- **Big Data**

MOBILE INVESTIGATIONS AND DIGITAL FORENSICS



The amount of data generated **in two days** is as much as all data generated **in human history until 2003.**

The **advancements in mobile technology** in combination with the **acceptance and widespread adoption of mobile devices** by the community have led to a significant rise in mobile forensics cases.

The digital forensics market is expected to grow **from $4.62B in 2017 to $9.68B by 2022**, an annual compound growth rate of almost **16%.**

MOBILE INVESTIGATIONS AND DIGITAL FORENSICS

Mobile devices receive data from many sources, such as:

- computers,
- cloud servers,
- social media platforms,
- network components,
- drones,
- smart vehicles,
- wireless cameras
- smart home devices,

while new technologies come into existence and are integrated into this diverse ecosystem with the progression of science and industry.

CLOUD COMPUTING

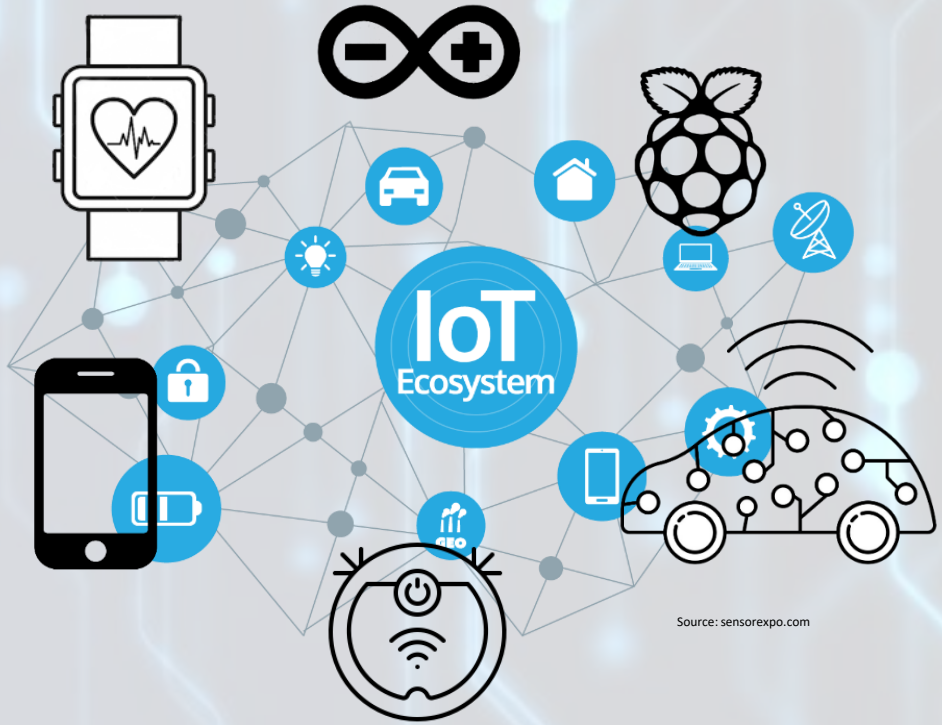
Cloud computing provides **large amounts of data** that can be utilized by examiners **to discover valuable artifacts** for criminal cases.

Cloud Computing Drivers:

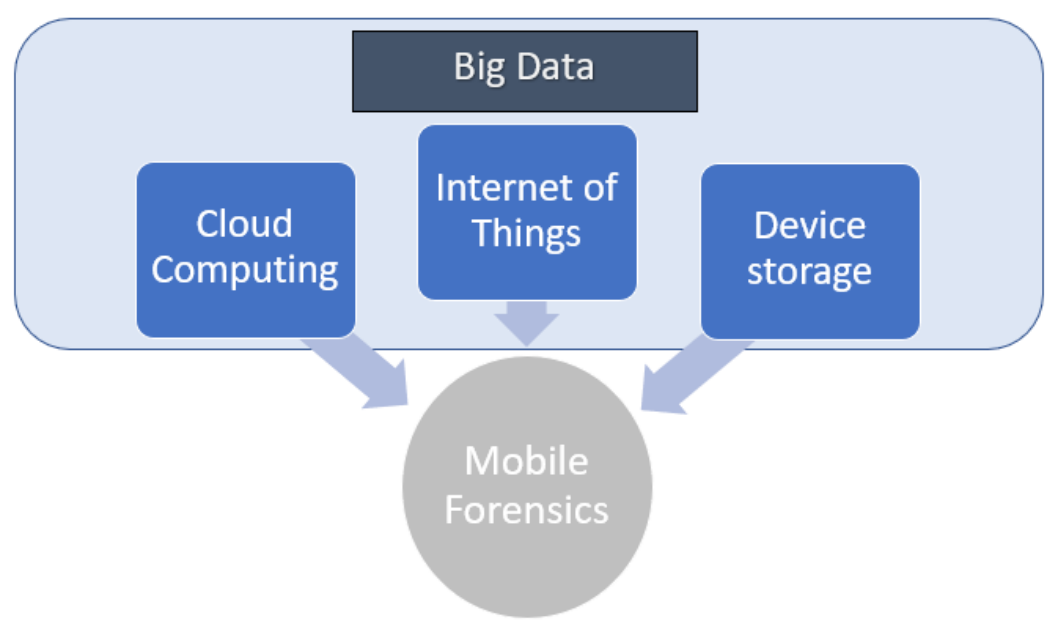
- ever-increasing **need for data storage services**,
- the availability of **high-capacity networks**,
- **low-cost computers and storage devices**,
- the widespread **adoption of hardware virtualization**,
- **service-oriented architecture** and
- **autonomic and utility computing.**

INTERNET OF THINGS

- More and more devices are part of daily **IoT Ecosystem** (e.g. IP cameras, vehicles, electric scooters, smart vacuum cleaners).
  - Gaining particular attention from cybercriminals.
  - May contain information of great forensic value.
  - IoT environment is deeply interconnected with smartphones and the respective applications.
- **Device storage** is being increased, hence the **data** to be examined.



BIG DATA



In a study done by IDC, it is expected that by 2025 we will have more than **175 zettabytes of data**.

Due to the conjunction of Big Data with **Information Technology**, **Cloud computing** and the **IoT ecosystem**, it is a particularly important research subject in the domain of Digital Forensics.

Big datasets come with **algorithmic challenges** that previously did not exist. Hence, there is seen by some to be a need to **fundamentally change the processing ways**.

From the analysis of the current status of Mobile Forensics, we can conclude that it is affiliated with all aspects of Digital Forensics, most notably with Cloud Forensics, IoT and Big Data Forensics.

ISSUES IN MOBILE FORENSICS

A. Volume of data

B. Variety and Variability of data

VOLUME OF DATA

Big Data:

- **high-volume**,
- **high-velocity** and/or
- **high-variety** information

Big Data demands **cost-effective**, **innovative forms of information processing** that enable **enhanced insight, decision making, and process automation (Gartner)**.

In mobile forensics, the meaning and value of the data volume to be examined and analyzed is inherently interconnected to the requirement for **fast, efficient procedures and techniques**, as well as **accurate and concrete results.**

VARIETY AND VARIABILITY OF DATA

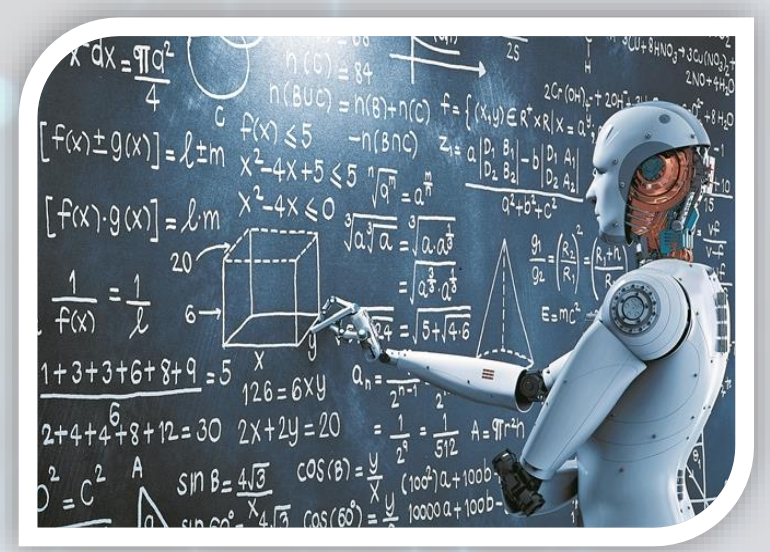
- ***Structured Data (****numbers, dates, groups of words or strings****)***
- ***Unstructured Data (****information that either does not have a pre-defined data model or cannot be structured in an orderly fashion, such as in ordered rows and columns as found in databases****)***

One of the main contributors to the variety and variability of data in mobile forensic investigations is the **IoT ecosystem**.

The most important challenges for Forensic Examiners:

- **data storage**,
- **data format**,
- **the diversity of IoT devices**,
- **support for these devices by current digital evidence software**.

MOBILE FORENSICS AND AI SOLUTIONS

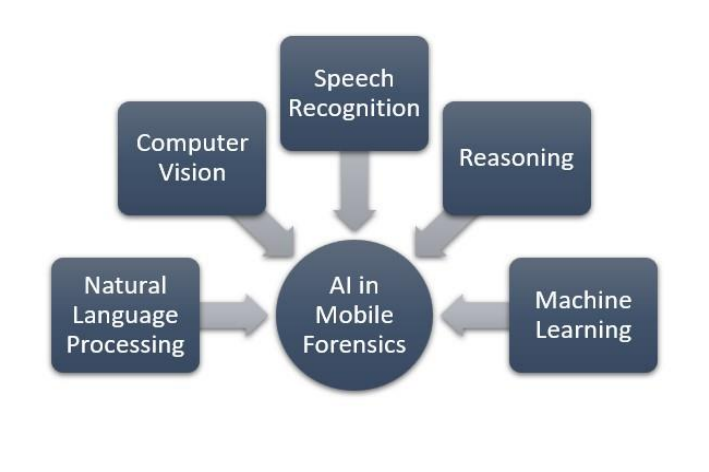


Artificial Intelligence has the potential for **providing the necessary expertise** and **helps in the standardization, management and exchange of a large amount of data, information and knowledge** in the forensic domain.

Results from of AI research:

- **reduction in the volume of evidence** to be examined
- **reduction in execution times** obtained with the distributed processing of the evidence
- implementation of systems that can **reduce human knowledge into a set of standardized rules**

CURRENT TOOLS

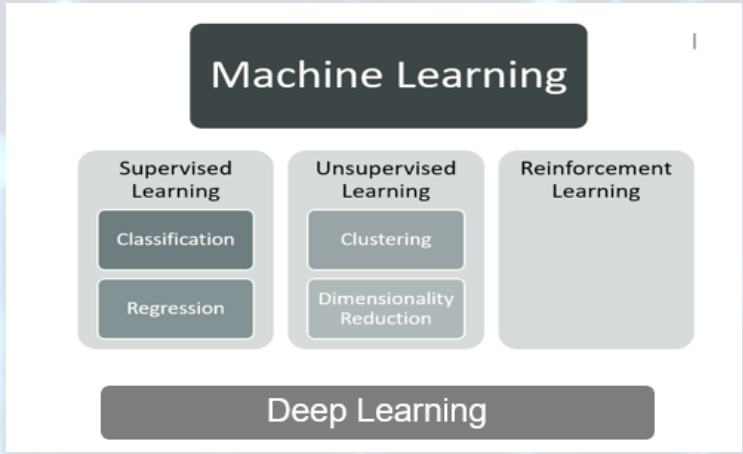


AI Techniques in Mobile Forensics:

- **Case Based Reasoners (CBRs),**
- **Pattern Recognition,**
- **Knowledge Discovery,**
- **System Adaptation,**
- **Refinement of Knowledge and**
- **Machine Learning (Symbolic Learners and Sub Symbolic Learners)**

Machine Learning:

- ✓ **Supervised Learning**
- ✓ **Unsupervised Learning**
- ✓ **Reinforcement Learning**
- ✓ **Deep Learning - Artificial Neural Networks**



CURRENT TOOLS

Important **forensic processes** by current tools:

- **image classification** and
- **video classification**

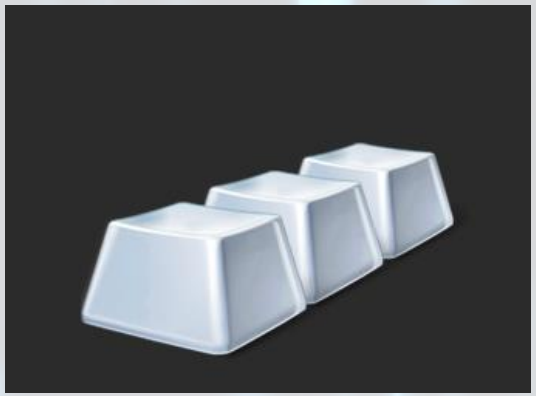
for predetermined categories, such as weapons, documents, nudity, faces and vehicles.

Other **useful AI functions**:

- Speech-To-Text Recognition
- Chat Classification
- Text Document Analysis - Text Clustering

Improvements into the accuracy of image and video classification by forensic tools would be **very beneficial** for the field of Mobile Forensics.

Examples of **image misclassification** into the category "Nudity":



TECHNOLOGY INTEGRATION AND LEGAL ISSUES

Challenges for the complete integration of AI related forensic technologies:

- lack of proper **regulatory framework**
- the general **fear and absence of trust** for the technology
- **shortage of computer systems** capable of supporting AI applications and features
- **complexity** in AI and ML algorithms
- **insufficiency of relevant datasets**, which are necessary for machine learning

TECHNOLOGY INTEGRATION AND LEGAL ISSUES

Legal Issues for the complete integration of AI related forensic technologies:

- **Legal Value of Artifacts**
- **Algorithmic Transparency**
- **Protection of Data Privacy**

Assessment of Evaluation metrics:

- **Accuracy**
- **True Positive Rate (Sensitivity)**
- **True Negative Rate (Specificity)**
- **Precision**
- **Recall**
- **F1-Score**

CONCLUSIONS

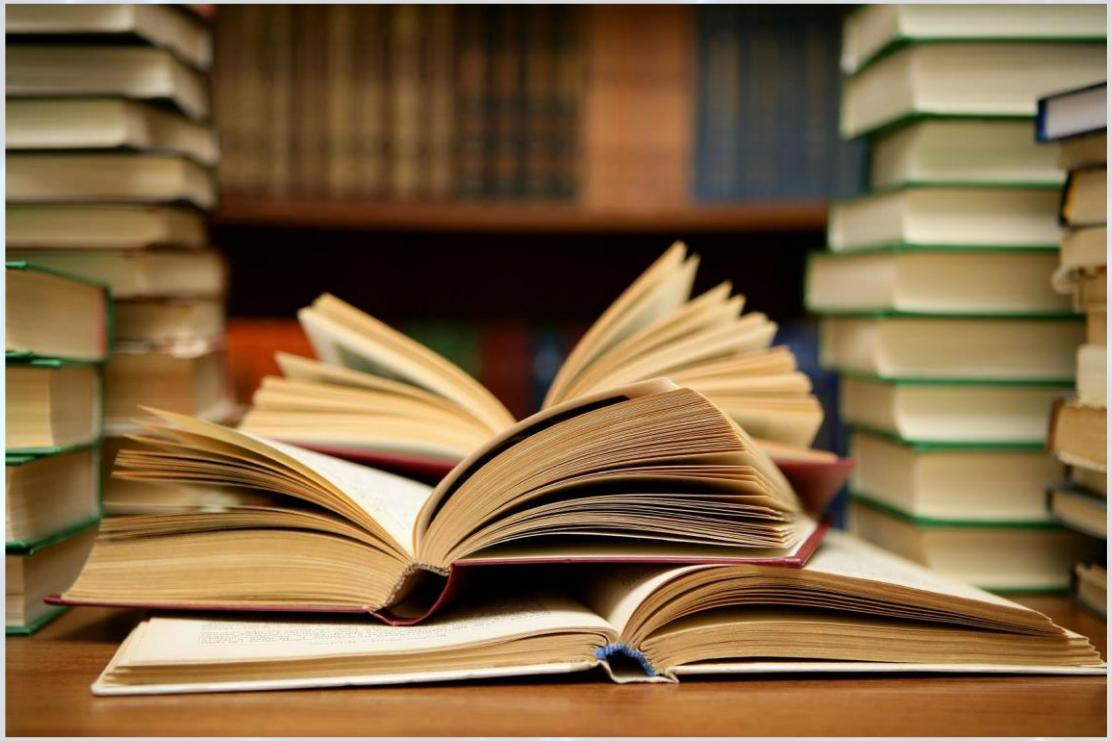
The establishment and development of **AI and Machine Learning techniques** in the field of Mobile Forensics seems inevitable and it could **revolutionize the practice of digital forensics investigations**.

Legal issues should be taken into consideration and regulatory measures should be put in place, in order to **utilize the benefits of the research with respect to legal proceedings and data privacy**.

The rapid advancements in **Mobile Networks, Cloud Computing, Internet of Things and Big Data technologies** indicate that a new era in Digital and Mobile Forensics is emerging, so **the potential and the concerns regarding Artificial Intelligence should be examined as soon as possible**.

RECOMMENDATIONS FOR FUTURE WORK

A. **Pattern Recognition – Computer Vision**
B. **Natural Language Processing**
C. **Open – source tools**
D. **Standardized procedures and legal issues**



REFERENCES

- Market Insider, Digital Forensics Market - Global Forecast to 2022 (16 March 2018) Retrieved from "https://markets.businessinsider.com/news/stocks/digital-forensics-market-global-forecast-to-2022-1018885400"
- Homem, I. (2018). Advancing Automation in Digital Forensic Investigations. *Academic dissertation for the Degree of Doctor of Philosophy in Computer and Systems Sciences at Stockholm University to be publicly defended, December 2018.*
- Mohammed, H.J., Clarke, N., & Li, F. (2016). An Automated Approach for Digital Forensic Analysis of Heterogeneous Big Data. J. Digit. Forensics Secur. Law, 11, 137-152.
- Jarrett A, Choo K-KR. The impact of automation and artificial intelligence on digital forensics. WIREs Forensic Sci. 2021;e1418. https://doi.org/10.1002/wfs2.1418
- Harrill, D.C., &Mislan, R.P. (2007). A Small Scale Digital Device Forensics ontology.
- Roy, Nihar& Khanna, Anshul &Aneja, Leesha. (2016). Android phone forensic: Tools and techniques. 605-610. 10.1109/CCAA.2016.7813792.
- https://www.ibm.com/cloud/learn/what-is-mobile-cloud-computing".
- Khan, A. u R.; Othman, M.; Madani, S. A.; Khan, S. U. (2014-01-01). "A Survey of Mobile Cloud Computing Application Models". IEEE Communications Surveys and Tutorials.
- Data Age 2025, sponsored by Seagate with data from IDC Global DataSphere, Nov 2018. IDC White Paper.
- https://medium.com/digital-reflections/cloud-computing-101-2b66e54c66c4.
- https://www.forbes.com/sites/louiscolumbus/2014/03/14/roundup-of-cloud-computing-forecasts-and-market-estimates-2014/?sh=1746969757a2.
- NIST, The NIST Definition of Cloud Computing, SP800-145, September 2011.
- "https://en.wikipedia.org/wiki/Cloud_computing#Mobile_%22backend%22_as_a_service_(MBaaS)".
- "https://www.programmableweb.com/news/builtio-building-enterprise-mbaas-platform-iot/interview/2014/03/03".
- "https://www.oracle.com/internet-of-things/what-is-iot/".
- eForensics Magazine, "IoT Forensics", VOL.08 N0.06, Issue 06/2019, (91) June, ISSN 2300 6986, page 37.
- Nordrum, Amy (18 August 2016). "Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated". IEEE Spectrum.
- E. Sejdić (March 2014). "Adapt current tools for use with big data". Nature. 507 (7492): 306.
- Alzaabi M, editor Ontology-based forensic analysis of mobile devices. Electronics, Circuits, and Systems (ICECS), 2013 IEEE 20th International Conference on; 2013: IEEE.
- "https://www.gartner.com/en/information-technology/glossary/big-data".
- Gantz, J. and E. Reinsel. 2011. "Extracting Value from Chaos", IDC's Digital Universe Study, sponsored by EMC.
- R. Quick, K.-K.R. Choo, Big Digital Forensic Data, Volume 2: Quick Analysis for Evidence and Intelligence. Springer Briefs on Cyber Security Systems and Networks, Springer (2018)
- Miorandi, D., Sicari, S., De Pellegrini, F., & Chlamtac, I. (2012). Internet of things: Vision, applications and research challenges. Ad Hoc Networks, 10(7), 1497–1516. doi:10.1016/j.adhoc.2012.02.016
- Ibrar Yaqoob, Ibrahim Abaker Targio Hashem, Arif Ahmed, S.M. Ahsan Kazmi, Choong Seon Hong, Internet of things forensics: Recent advances, taxonomy, requirements, and open challenges, Future Generation Computer Systems, Volume 92, 2019, ISSN 0167-739X, https://doi.org/10.1016/j.future.2018.09.058.

REFERENCES

- https://en.wikipedia.org/wiki/Digital_forensics. Retrieved 02 September 2021.
- Hoelz, Bruno & Ralha, Célia & Geeverghese, Rajiv & Junior, Hugo. (2008). A Cooperative Multi-agent Approach to Computer Forensics. Proceedings - 2008 IEEE/WIC/ACM International Conference on Intelligent Agent Technology, IAT 2008. 477-483. 10.1109/WIIAT.2008.55.
- C. Platzer, M. Stuetz, and M. Lindorfer, "Skin sheriff: a machine learning solution for detecting explicit images," in Proceedings of the 2nd international workshop on Security and forensics in communication systems, 2014, pp. 45-56: ACM
- Poole, D., Mackworth, A., & Goebel, R. (1998). Computational intelligence: A logical approach. Oxford: Oxford University Press.
- Russell, S. J., &Norvig, P. (2010). Artificial intelligence: A modern approach. Applied Mechanics & Materials, 263(5), 2829–2833.
- Stefania Costantini, Francesca A. Lisi, Raffaele Olivieri. Knowledge Representation and Reasoning meets Digital Forensics: The COST Action DigForASP (short paper). RCRA/RiCeRcA@AI*IA. January 2019.
- https://www.samsung.com/semiconductor/minisite/exynos/technology/ai/".
- Faye Mitchell,The use of Artificial Intelligence in digital forensics: An Introduction, 2010.
- Rowena Rodrigues, Legal and human rights issues of AI: Gaps, challenges and vulnerabilities, Journal of Responsible Technology, Volume 4, 2020, 100005, ISSN 2666-6596, https://doi.org/10.1016/j.jrt.2020.100005.
- Qadir, A. M., &Varol, A. (2020). The Role of Machine Learning in Digital Forensics. 2020 8th International Symposium on Digital Forensics and Security (ISDFS). doi:10.1109/isdfs49300.2020.9116298
- J. Mena, Investigative data mining for security and criminal detection. Butterworth-Heinemann, 2003.
- J. Mena, Machine learning forensics for law enforcement, security, and intelligence. Auerbach Publications, 2016.

Thank you for your time!

QUESTIONS?

n.papadoudis@astynomia.gr

CEPOL 79/2022 European Research and Science Conference
8-10 June 2022, MRU, Vilnius, Lithuania