

Integrovaná platforma pro zpracování digitálních dat z bezpečnostních incidentů (TARZAN)

Identifikační kód VI20172020062

Název předkládaného výsledku: *Workshop on Evidence Collection and Analysis of Webpages*

Typ výsledku dle UV č. 837/2017	Evidenční číslo (příjemce)	Rok vzniku
A audiovizuální tvorba		2019
ISBN-ISSN	Webový odkaz na výsledek	Kde a kdy publikováno
	https://www.fit.vut.cz/research/publication/12148/	ISS World Asia 2019

Anotace k výsledku:

Kybernetické zločiny, jako je ransomware, kyberšikana, podvod, nezákonné činnosti temných sítí, nesprávné šíření sexuálního obsahu nebo dokonce phishing, jsou velmi nestabilní, pokud jde o shromažďování důkazů. Webové stránky související s těmito zločiny jsou obvykle k dispozici pouze na pár dní, někdy i hodin. Workshop představuje metody, jak tyto webové stránky efektivně stahovat, dekodovat, analyzovat a archivovat. Zaměřuje se na bezpečný a kontrolovatelný sběr cenných (meta) dat, která lze později použít jako důkaz. Prezentace nastiňuje teorii moderního webdizajnu (HTML, CSS, Java / TypeScript), známé knihovny pro škrabání a dekodování (např. Scrapy, Selenium) a aktuální výzvy (například jednostránkové aplikace, přístup k dynamickému obsahu), spuštění JavaScriptu). Relace zahrnuje ukázky procesu sběru a existujících nástrojů. Účastníci obdrží náš nástroj s otevřeným zdrojovým kódem, který snadno archivuje daný obsah URL spolu se základní sadou metadat.

Řešitelský tým: Petr Matoušek (manažer a hlavní řešitel), Vladimír Veselý (realizační tým)