Design, architecture, and technology used for the development of the LOCARD platform

One of the most attractive outcomes of the LOCARD project is the holistic forensic platform where law enforcement agencies, forensic laboratories and judicial authorities, among others, can use to preserve the integrity of the chain of custody of digital evidences collected from criminal activities. The LOCARD's public deliverable D3.5 (see **HERE**) describes the LOCARD architecture, its components, and main technologies. To this end, implementation details are provided at a modular level, by describing the functionality and operation of each module, as well as at an integration level between the different modules.

The high-level description of the LOCARD architecture, including the conceptual and logical of the whole architecture and each individual module, is described in Section **3**. To put the architecture in context, Section **3.1.** provides a conceptual overview of the LOCARD system, by defining the role of each entity within the system at a high level. The logical architecture of LOCARD is elaborated in Section **3.2.**, by describing a multi-tier web-based architecture composed of multiple basic tiers, such as the client (or presentation) tier that interacts with the end users, the application tier comprising the business logic, the integration layer for connecting applications and enabling data exchange, and the data tier considering several databases and data repositories. J2EE, React, NodeJS, RabbitMQ and MariaDB are just some of the main technologies used to develop the LOCARD platform. Complementing the aforementioned tiers, the architecture also considers both security and management tiers to support authentication and supervision tasks. The main LOCARD blocks, resulting from the clear modularity of the LOCARD design, are presented in Section **3.3.** In a nutshell, the LOCARD Portal, containing critical system's modules, such as the identity management, the storage management, and the blockchain are highlighted. A more comprehensive description of the LOCARD forensic flow is then detailed in Section **3.4.** following standard practices like the ones recommended by Interpol. Moreover, the physical architecture used in the LOCARD development to achieve high availability, performance and security is proposed in Section **3.5.**

At a more detailed level, Section **4** provides readers with a full detailed description of the functional, technical and interconnection details of all the modules implemented in the LOCARD system. Hereafter, a summary of the different modules that readers can find in the public deliverable is provided. First of all, Section **4.1.** describes the intelligent crawler used

LOCARD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement nº 832735.

to facilitate the collection of digital evidences from the Internet by means of the already existing HTTrack software technology. Mainly oriented to the investigators' related tasks, Section **4.2.** explains the investigator's toolkit allowing them to perform their usual data acquisition and analysis procedures likewise with any current forensic tool. The communication engine implemented in LOCARD, detailed in Section **4.3.**, allows properly authenticated end users to communication between them so to request further details of a specific evidence, report or case. As citizens are also able to report illicit content (abuse over a public feed, illegal streaming content, etc.), Section **4.4.** addresses the specifics of how LOCARD allows these users to fulfil this task through its crowdsource intelligence module. Another paramount module is the storage manager, detailed in Section **4.5.**, responsible for storing different types of data (e.g., evidences, metadata…) in the many LOCARD databases. The module in charge of managing the federated permissioned LOCARD blockchain network and issuing the smart contracts that enable end user's interactions is described in Section **4.6.** Especially in LOCARD, authentication is a crucial feature implemented in the user and identity manager module, described in Section **4.7**, responsible for identifying and authorizing the access of end users, trusted third parties and jurisdictional entities to the LOCARD platform. To this end, a custom OAuth2 service defining access, right, roles and segregation of data for every user is defined.

The collaboration between different end users and the enhancement of criminal investigations are two key features of LOCARD. In this regard, the intelligence engine, explained in Section **4.8.**, provides intelligence to extract knowledge and correlations to ease criminal investigations. The processing of similarly structured data by means of semantic analysis algorithms plays a major role to achieve successful cross-correlations. Another added-value feature of LOCARD is the deviant patterns repository, highlighted in Section **4.9.** The task of this module is twofold. On the one hand, it monitors social live streaming services and social media to identify deviant behaviours related to criminal activity (e.g., child exploitation, predatory acts, etc.) using machine learning, natural language processing and social network analysis. On the other hand, the module must define which is the ground truth of a deviant behaviour in LOCARD, by considering curated datasets as well as trained machine learning based models used for the detection of online deviant behaviours. Section **4.10.** briefly describes the LOCARD connector used to link the proper LOCARD platform with external repositories, with the aim to increase the visibility of LOCARD with external data repositories for possible collaborations. With a focus on the forensic outcome, Section **4.11.** explains the reporting engine responsible for collecting the data related to a specific case and creating an audit log of all the interactions related to it. The resulting report will contain all the interactions logged into the blockchain, such as the user identifier reporting. Similarly to most applications, LOCARD also considers an alert engine, presented in Section **4.12.**, that notifies users upon certain predefined conditions: e.g., when a new interaction is stored in the blockchain for a specific investigation, all the investigation-related users will receive a message/alert, specifying that a new event has been logged. This alert module helps to reduce investigation times and enhance fight response. With the aim to achieve secure storage, cryptographical operations and secure external execution, LOCARD considers a trusted execution environment technology, explained in Section **4.13.**, to isolate specific and extremely sensitive operations. Another LOCARD core module is the core orchestrator bus, specified in Section **4.14.**, enabling the communication between the different modules considered in LOCARD. The latest module, detailed in Section **4.15.**, is the LOCARD Portal, i.e., the front-end application through which all the modules of the platform will be accessed.

LOCARD project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement nº 832735.